# Characterisation of the 5' flanking region of the murine PrP gene.

John Daniel Cameron.

PhD. Molecular biology

University of Edinburgh 1995.

## Dedication

This thesis is dedicated to my Mum and sister, Fiona, for their support through the years.

# Acknowledgements

Firstly, I would like to thank my supervisor, Dr Jean Manson, for her enthusiasm in my project and for our frank discussions.

Thanks, too, to Herbert Baybutt and Carol Gilchrist for their help in the lab. Special thanks to Wilfred Goldmann for his approachable attitude and constant encouragement. Also, I appreciate the comments from Wilfred, Jim Hope and Mary Bownes with regard to reading this thesis.

I am also grateful for the PrP cosmid clone which David Westaway made available.

Finally, a special thanks to the youngsters, and the not so young, at the unit who made my stay there a very enjoyable one.

## Abstract

Scrapie, Creutzfeldt-Jakob disease and bovine spongiform encephalopathy belong to a group of diseases known as the transmissible degenerative encephalopathies. These diseases are characterised by astrocytic gliosis, neuronal vacuolation, deposition of amyloid plaques, long incubation periods and by their transmissibility to laboratory rodents.

PrP has a central role in the pathogenesis of the transmissible degenerative encephalopathies. During disease PrP accumulates pathologically in and around cells of the brain. In uninfected animals PrP is a membrane bound sialoglycoprotein of molecular weight 33-35 kDa. PrP mRNA is found in most tissues, but at differing levels. The brain expresses PrP mRNA highly, whereas organs such as heart and kidney express PrP mRNA to a lesser extent. Liver and spleen express PrP at very low levels. The expression of PrP mRNA has also been shown to be regulated during mouse development, both postnatally and embryonically. PrP levels rise four-fold from day 0 to day 20 and are then maintained at this level throughout adult life. Prenatally, PrP has been detected by in situ hybridisation at 13.5 days in the developing brain and spinal cord and by 16.5 days PrP transcripts could be detected also in specific non-neuronal populations. Nerve growth factor has been shown to influence the levels of PrP *in vivo* and *in vitro*. Injection into hamster brain increased PrP mRNA levels approximately nine-fold and PrP protein levels three-fold.
The regulatory processes involved in controlling these types of expression are not known.

This project sets out to examine the transcriptional regulation of the mouse PrP gene by studying the promoter region and the first intron.

This was done by isolating a 3.5 kb BamHI / KpnI fragment from a murine PrP cosmid clone. This subclone contained 1.2 kb of 5' region to exon 1. This exon was shown to be separated from exon 2 by a 2.3 kb intron.

Sequence analysis of the 5' flanking region showed that it was highly G+C rich, contained no TATA box and had potential binding sites for many transcription factors including Sp1, AP1, AP2 and EGR1.

In a transient transfection assay this 3.5 kb BamHI / KpnI fragment was shown to act as a promoter by linking it to the bacterial reporter gene - chloramphenicol acetyl transferase and introducing it into neuro2a cells. Subsequent deletions both in the 5' flanking region and in the intron revealed positive regulatory elements in both.

The transcriptional start sites of the PrP gene were mapped using the primer extension technique with an oligonucleotide whose target sequence resides in exon 2.

These reporter gene constructs will be used to generate transgenic mice and to assess their role, more fully, in the regulation of the PrP gene.

# Abbreviations

| | |
|---|---|
| A | Adenine |
| Amp | Ampicillin |
| B | G or C or T |
| β-gal | Beta-galactosidase |
| bp | Base pair |
| BSA | Bovine serum albumin |
| BSE | Bovine spongiform encephalopathy |
| $^{14}C$ | Carbon-14 |
| C | Cytosine |
| CAT | Chloramphenicol acetyl transferase |
| Ci | Curie (s) |
| CIP | Calf intestinal phosphatase |
| CJD | Creutzfeldt-Jakob disease |
| CoA | Coenzyme A |
| CPM | Counts per minute |
| CWD | Chronic wasting disease |
| D | G or A or T |
| ddNTP | Dideoxynucleoside 5'-triphosphate |
| DEPC | Diethyl pyrocarbonate |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTP | Deoxynucleoside 5'-triphosphate |
| DMEM | Dulbeccos modified Eagles medium |
| DMSO | Dimethylsulphoxide |
| DTT | Dithiothreithol |
| EDTA | Diaminoethanetetra-acetic acid disodium salt |
| FCS | Foetal calf serum |
| FFI | Fatal familial insomnia |
| g | Relative centrifugal field |
| G | Guanine |

| | |
|---|---|
| GABA | Gamma aminobutyric acid |
| GFAP | Glial fibrillary acidic protein |
| GSS | Gerstmann-Straussler-Scheinker syndrome |
| H | A or T or C |
| Hr | Hour (s) |
| K | G or T |
| kb | Kilobase pair |
| kDa | Kilodalton |
| *lac* Z | β-galactosidase gene |
| LMP | Low melting point |
| M | A or C |
| Min | Minute (s) |
| mRNA | Messenger RNA |
| N-CAMs | Neural cell adhesion molecules |
| NGF | Nerve growth factor |
| nm | Nanometres |
| OD | Optical density |
| ONPG | o-nitrophenyl-β-D-galactopyraniside |
| ORF | Open reading frame |
| $^{32}P$ | Phosphourus-32 |
| PCR | Polymerase chain reaction |
| pfu | Plaque forming units |
| PrP | Prion protein |
| R | A or G |
| RNA | Ribonucleic acid |
| RPM | Revolutions per min |
| R/T | Room temperature |
| RT-PCR | Reverse transcription- PCR |
| $^{35}S$ | Sulphur-35 |
| S | G or C |
| SD | Standard deviation |
| SDS | Sodium dodecyl sulphate |

| | |
|---|---|
| Sec | Second (s) |
| SEM | Standard error of the mean |
| T | Thymine |
| TDE | Transmissible degenerative encephalopathy |
| T.E | Tris-EDTA (pH 8) |
| UTR | Untranslated region |
| UV | Ultraviolet |
| W | A or T |

Table of contents

# Chapter 1 Introduction

## 1.1 Transmissible degenerative encephalopathies

The transmissible degenerative encephalopathies (TDEs) are a group of naturally occurring and fatal, neurological, diseases of both humans and animals.
Scrapie is the best studied of this group, its existence having being documented centuries ago, and occurs naturally in sheep and to a lesser extent in goats. The main mode of transmission in natural scrapie is thought to be maternal. However, horizontal transmission of natural scrapie also seems to occur (Dickinson *et al.*, 1974). Rodents, especially mice, are frequently used as experimental models of the disease. Bovine spongiform encephalopathy (BSE) first appeared in cattle in 1986 (Wells *et al.*, 1987) and epidemiological evidence later suggested that infection occurred by ingestion of contaminated food-stuffs (Wilesmith *et al.*, 1991). This coincided with changes in the way which animal waste tissues were rendered. These changes meant that scrapie agent could now survive the rendering process and was therefore able to enter the food chain of animals which consumed it. In addition to cattle, scrapie-like diseases soon became manifest in exotic zoo animals such as nyala, gemsbok and kudu and these, too, were thought to be due to consumption of contaminated food-stuffs. Chronic wasting disease (CWD) affects mule deer and elk whereas transmissible mink encephalopathy, as the name suggests, affects mink. Both of these diseases are thought to stem from ingestion of scrapie infected tissues.

In humans there are four diseases which can be classified into the TDEs. These are Creutzfeldt-Jakob disease (CJD), Gerstmann-Straussler-Scheinker syndrome (GSS), Kuru and the recently discovered fatal familial insomnia (FFI). CJD was first characterised at the turn of the century and since then has been shown to exist in three forms: sporadic, familial and iatrogenic. The majority of CJD cases are of a sporadic nature in that neither a genetic (familial) nor an accidental (iatrogenic) route of transmission can be found in such cases. The transmissibility of sporadic CJD is not understood. Unlike CJD, most cases of GSS are familial in nature, characterised

by an autosomal pattern of inheritance. The majority of patients with GSS have substitution mutations in the open reading frame of their PrP gene (for review see Hsiao and Prusiner, 1990). Kuru occurs only in the Fore tribe in Papua New Guinea but its prevalence is now declining since its route of transmission was established. Ritualistic cannibalism was traced as this route of transmission (Gajdusek, 1977). The TDEs are characterised by long incubation periods, transmissibility to laboratory animals, astrocytic gliosis and by neuropathology such as spongiform degeneration of neurones and deposition of amyloid plaques. The precise nature of this neuropathology can be quite variable and depends on the disease type, the species of animal and its genotype, and strain of agent (defined only in rodents).

PrP protein has a critical role in the pathogenesis of these diseases. In uninfected animals PrP is found on the cell surface of many cell types, but especially neurones. During disease, however, PrP protein becomes altered in some undefined way. These alterations are poorly understood, but their consequences are that PrP aggregates pathologically in and around cells of the brain.

In order to understand these diseases it is first necessary to understand at what level the regulation of PrP is affected. Several workers have suggested that there may be a post-translational (Borchelt *et al.*, 1990) or conformational (McKinley *et al.*, 1991) change in PrP during disease and that this causes infection-specific pathology to form and, indeed, is the basis of infectivity itself. Despite much effort, the exact nature of this change has not been determined (Stahl *et al.*, 1993). There remains, however, the possibility that the regulation of PrP is affected, during disease, at a transcriptional or post-transcriptional level. These possibilities must initially be addressed by examining the regulation of PrP gene expression in uninfected animals. By doing this one can start to determine the role of such events during disease.

## 1.2 PrP Protein

PrP (prion protein) is a cellular sialoglycoprotein (Bolton *et al.*, 1985) which is attached to the outside of a variety of cells by a glycophosatidylinositol anchor (Stahl *et al.*, 1987). PrP can exist as two different isoforms: the cellular form ($PrP^{c}$) which is found in healthy animals and the scrapie isoform ($PrP^{sc}$) which is found only in diseased animals. $PrP^{c}$ is heterogenous in size and has a molecular weight ranging from 33-35 kDa (Turk *et al.*, 1988; Bolton *et al.*, 1987). $PrP^{c}$ is soluble and can be digested to completion with proteinase k (Meyer *et al.*, 1986; Bendheim *et al.*, 1988). $PrP^{sc}$, on the other hand, is specific to the transmissible degenerative encephalopathies and has been found to be insoluble and partially resistant to proteinase k treatment, resulting in a resistant core of 27-30 kDa. This protease resistance is characteristic of the TDEs.

Both isoforms of PrP have been shown to shown to undergo the same extensive post-translational modifications; cleavage of the N-terminal signal peptide (Basler *et al.*, 1986; Hope *et al.*, 1986; Turk *et al.*, 1988), N-linked glycosylation at their predicted sites (Caughey *et al.*, 1988) and addition of a glycophosphatidylinositol anchor at its C-terminus after cleavage of its terminal peptide (Turk *et al.*, 1988., Stahl *et al.*, 1990). The difference between $PrP^{c}$ and $PrP^{sc}$ is not known, however similarities in size and charge suggest that this difference is slight and may reflect either a conformational or post-translational modification (Hope *et al.*, 1986; Turk *et al.*, 1988).

Based on DNA sequences, both cDNA and genomic, of PrP genes from various species the deduced amino-acid sequences have been shown to have a greater than 85 % identity with each other (for review see Goldmann, 1993). This homology of PrP proteins has been shown to occur in mouse (Locht *et al.*, 1986., Westaway *et al.*, 1987), hamster (Basler *et al.*, 1986), human (Kretsczmar *et al.*, 1986) and sheep

(Goldmann *et al.*, 1990). This high degree of conservation suggests that PrP has an important function. This function, however, has not been determined, despite several speculations.

During lymphocyte activation PrP protein increases in abundance on the surface of lymphocytes. Cashman *et al* (1990) showed that antibodies to PrP can block this activation. This led them to suggest that PrP, in lymphocytes at least, may play a role in signal transduction. Others have suggested that because a GFAP (glial fibrillary acidic protein) stimulating factor is produced from a cell line derived from a CJD infected brain that PrP may act in controlling astrocyte proliferation (Olescak *et al.*, 1989). Due to its location on the external surface of cells, PrP has also been suggested to perform a similar function as the N-CAMs and Cadhedrins (Hope and Manson, 1991). That is, in directing and maintaining the architecture of the nervous system. In addition, PrP may play a role in embryogenesis (Manson *et al.*, 1992 a).

To determine what physiological role PrP plays and to understand the relationship which PrP has with the causative agent, several labs have generated mice in which the PrP genes have been ablated or "knocked-out" (Bueler *et al.*, 1992; Manson *et al.*, 1994). These are called PrP null mice.

Bueler *et al* (1992) generated PrP null mice by disrupting one allele of the PrP gene via a homologous recombination event. This involved replacing codons 4-187 of the PrP open reading frame with a neomycin phosphotransferase gene under the control of the HSV thymidine kinase promoter. Cells which contained the disrupted PrP gene were injected into 4 day old blastocyts and implanted into foster mothers. Heterozygotes were mated with each other to produce mice which were homozygous for the disrupted PrP gene. These $PrP^{0/0}$ null mice showed normal development and had no abnormalities up to the age of seven months. Disruption of the PrP gene produced no gross abnormalities at both macroscopic and microscopic level in the brain and skeletal muscle (Bueler *et al.*, 1992). These null mice showed no significant difference in the response of splenocytes, when compared to $PrP^{+/+}$ mice, to activation by concanavalin A. In addition, PrP null mice showed no differences in learning ability, based on 3 simple learning tests and showed no breeding

differences when compared to wild type mice. Recently, however, a phenotypic difference has been found between wild type mice and null mice (Collinge *et al.*, 1994). PrP null mice have an apparently weakened $GABA_A$ ($\gamma$-aminobutyric acid type A) receptor-mediated fast inhibition response and also impaired long term potentiation. This defective synaptic inhibition was detected in the hippocampus. Collinge *et al* (1994) argue that during disease, $PrP^c$ loses its function (presumably by being converted to $PrP^{sc}$) and that this loss of function contributes to early synaptic loss (Clinton *et al.*, 1993) and to neuronal degeneration.

### 1.3 The scrapie agent: its properties and relationship to PrP

One of the first indications that the causative agent of scrapie was different from other infectious organisms came from the work of Alper *et al* (1967, 1978) who used ionising and U.V irradiation in inactivation of infectivity studies. These experiments indicated that the putative genome of the agent must be very small due to its high resistance to inactivation by irradiation. The interpretation of this data was called into question by Rowher (1984) who proposed that the genome of the agent was similar to that of a virus both in terms of its size and by its susceptibility to inactivation by irradiation. However, the original interpretation of the data by Alper et al was substantiated by more recent experiments which support this (Bellinger-Kawahara *et al.*, 1987, 1988). Other experiments have shown that scrapie infectivity is not affected by the action of nucleases (Prusiner, 1982). Furthermore, to date no one has managed to isolate a scrapie specific nucleic acid from infectious tissue fractions despite numerous attempts (Wietgrefe *et al.*, 1985; Meyer *et al.*, 1991; Kellings *et al.*, 1992).

Much evidence has now amassed which strongly suggest that the agent is composed of protein and if this protein is destroyed or modified then this has a deleterious

effect on infectivity. By using procedures which purified infectivity from scrapie infected hamster brain McKinley *et al* (1983) showed that PrP was the predominant protein in fractions purified for scrapie infectivity. Furthermore, the concentration of PrP was found to be directly proportional to the amount of infectivity and that the isoform of PrP was $PrP^{sc}$ based on it's resistance to proteinase k. In 1982, Prusiner demonstrated using fractions purified for infectivity that if the proteins found in these fractions were treated with proteinase k, phenol, SDS, or diethylpyrocarbonate then infectivity was greatly diminished.

These lines of evidence lead to the proposal of the term "prion" to describe the infectious agent. This was introduced to distinguish it from viruses and viroids. Prions are defined as proteinaceous infectious particles which are resistant to inactivation by procedures which modify nucleic acids (Prusiner, 1982). Originally, prions were hypothesised to be either proteins which contained no genome, proteins which surrounded a nucleic acid (a virus) or proteins with which a small non coding, but perhaps regulatory, oligonucleotide was associated.

If any of these models are correct they must incorporate into them, and explain, the existence of distinct strains of scrapie. The existence of such strains is based entirely on the length of incubation period and neuropathological differences which they cause when injected into inbred lines of mice (Outram, 1976; Dickinson and Fraser, 1977). Like viruses, these scrapie strains can be mutated giving rise to other strains which have a different incubation period and which show different neuropathology in the same line of mouse. These changes are heritable and predictable and suggest that the scrapie agent contains its own independent genome (Bruce and Dickinson, 1987).

With a large degree of certainty all that can be said concerning the biochemical constitution of the agent is that it is composed or associated, at least in part, with the abnormal isoform of PrP.

## 1.4 Genetics of disease susceptibility

### 1.4.1 The *Sinc* and *Sip* genes

The incubation period can be defined as the time from intracerebral injection, of infected material, to onset of clinical symptoms. In mice, this incubation period can vary between 4 months and 2 years and is dependent on the strain of scrapie injected and on host genetic factors (Bruce *et al.*, 1991) . The *Sinc* gene (*s*crapie *inc*ubation period) is the major gene which controls incubation period in mice (Dickinson *et al.*, 1968) and there are two alleles of this gene: s7 and p7. In sheep the major gene which controls scrapie incubation period is the *Sip* gene and this too has two alleles: sA and pA. Evidence is now growing that *Sinc* (or *Sip*) is closely linked to PrP and in fact that the two may be congruent. This has been shown in mice (Carlson *et al.*, 1986., Hunter *et al.*, 1987) and in sheep (Hunter *et al.*, 1989) by using restriction fragment length polymorphisms (RFLP). By nucleotide sequencing Westaway *et al* (1987) have shown that a strain of mouse (NZW) which is s7 homozygous for *Sinc* differs from a mouse strain (I/Ln) which is p7 homozygous at predicted codons 108 and 189 of the PrP protein. This nucleotide transition, at predicted codon 108, would result in a leucine (NZW) to phenylalanine change (I/Ln). At codon 189 a change from valine (I/Ln) to threonine (NZW) was predicted. This therefore suggests that PrP and *Sinc* are the same gene but absolute proof is still lacking. If the genes are congruent it may be that these changes in the primary amino acid sequence of the PrP protein are responsible for the effects of incubation difference between mice with different Sinc genotypes.

Others have found that different genes may be involved in the disease process. Work with congenic mice which differ only in their major histocompatibility complex (H-2) suggested that another gene called Pid-1 (prion incubation determinant) plays a role in controlling incubation period (Kingsbury *et al.*, 1983). Pid-1 has been mapped close to the D region of the H-2 complex on mouse chromosome 17, but neither its function nor whether it actually represents one or more genes is known.

That Pid-1 maps to the D region of the murine H-2 region is interesting as class I MHC genes also map to this region and have been shown to increase their expression during the disease process. This increased expression has been shown by immunocytochemistry to be localised in neurones - the primary seat of destruction during disease (Duguid and Trzepacz, 1993). Beta 2 microglobulin is associated with the class I MHC gene product, to form the class I complex, and has also been shown to have increased expression during scrapie in hamsters (Duguid and Dinaver, 1990). Furthermore class II MHC genes were also shown to increase expression during disease (Duguid and Trzepacz, 1993). While these events may suggest that there is an attempted immune response to infection, there is no evidence of a detectable humoral or inflammatory response by the host (Gajdusek, 1977).

### 1.4.2 Mutations in the PrP gene which are linked to CJD and GSS

In animals such as mouse (Westaway *et al.*, 1987), hamster (Lowenstein *et al.*, 1990) and sheep (Goldmann *et al.*, 1990) differences in the nucleotide sequence, and deduced amino-acid sequence, of PrP have been shown to correlate with differences in scrapie incubation period.

In humans, familial CJD and GSS (Masters *et al.*, 1981) are linked to mutations in the open reading frame of the PrP gene (also known as the PRNP gene). There are however forms of these diseases, especially CJD, which have no apparent mutation in the open reading frame of the PrP gene and these are the sporadic and iatrogenic cases.

a) Gerstmann Straussler Scheinker syndrome

Most cases of GSS are familial, although some sporadic cases have been reported (for review see Hsiao and Prusiner, 1990). Out of the familial forms of GSS there are three basic types; each of which presents its own characteristic pathology and

clinical symptoms:

i) ataxic GSS

ii) dementing GSS

iii) GSS with neurofibrillary tangles

Despite the heritability of familial GSS, as characterised by an autosomal pattern of inheritance, it is also infectious as indeed are all other members of the transmissible degenerative encephalopathies.

GSS usually presents itself in the fourth decade of life and the average duration of the disease is five years, after which death ensues. GSS has a characteristically massive deposition of PrP amyloid and its incidence has been estimated at 1-10 per hundred million (see Hsiao and Prusiner, 1990).

The first and best characterised example of a mutation in the human PrP which linked to ataxic GSS was reported in 1989 by Hsiao and colleagues. They found that three individuals (two pedigrees) with GSS had a mutation at codon 102 of the PrP protein which changed a proline residue to a leucine residue. This mutation was not found in eight unaffected relatives or in one hundred unrelated normal individuals from the same ethnic background. Also, this mutation was not found in fifteen other individuals who had another form of transmissible degenerative encephalopathy. This suggested that although the 102 mutation may cause GSS, or at least play a role in the pathogenesis of disease, it was not responsible for any other disease within the TDE group (see section 1.6 for further analysis of the codon 102 mutation).

Subsequent analysis, by another group, showed that although the GSS 102 leucine →proline mutation is the most common mutation associated with GSS, it is not the only mutation. A codon 117 alanine →valine mutation can occur instead (Doh-ura *et al.*, 1989). This mutation is associated with dementing GSS.

The third form of GSS is unique in that all affected individuals have widespread Alzheimer-like neurofibrillary tangle (nft) deposition. These plaques immunolabel with anti-PrP antibodies but not with anti β-amyloid antibodies (Roberts *et al.*, 1986). GSS with neurofibrillary tangle deposition has been best characterised in two families; from Indiana and Sweden. Each of the affected members of the families had a mutation in the open reading frame of the PrP gene. The Indiana family had a mutation at codon 198 which changed a phenylalanine residue to a serine. The Swedish family, on the other hand, had a different mutation at codon 217 which substituted an arginine residue for a glutamine (Dlouhy *et al.*, 1992; Hsiao *et al.*, 1992).

**Table 1. Mutations associated with GSS**

| GSS form | | Mutation |
|---|---|---|
| Ataxic | | 102 proline → leucine |
| Dementing | | 117 alanine → valine |
| With NFT | Indiana | 198 phenylalanine → serine |
| | Swedish | 217 glutamine → arginine |

b) Creutzfeldt-Jakob disease

In comparison to GSS, CJD occurs with a relatively higher frequency affecting approximately 1 in a million people. 5 - 10 % of these cases are familial and are characterised by an autosomal dominant pattern of inheritance. The average age of death in patients with CJD is 51 years and the duration is usually shorter than GSS, lasting about 1 year. CJD is characterised by subacute dementia and widespread spongiform degeneration of neurones (Masters *et al.*, 1981).

Many studies have now linked several mutations in the human PrP gene to CJD. These mutations are both point and insertion mutations and a number of these will be described below.

The first point mutation in the PrP gene to be linked to CJD was reported by Goldfarb *et al* in 1990 (a). A mutation in codon 200 of the PrP which substituted a lysine residue for a glutamine was described initially for two clusters of CJD in Slovakia (Goldfarb *et al.*, 1990 a). The same mutation was subsequently found in individuals of different ethnic origin who were affected with CJD (Goldfarb *et al.*, 1990 b). A codon 178 mutation has also been reported which causes an asparagine to aspartic acid change (Brown *et al.*, 1992). As well as linking to CJD, this mutation has been linked to a recently discovered disease called fatal familial insomnia, FFI, (Medori *et al.*, 1992). This disease is characterised by sleep disturbances, ataxia and memory impairment. Neurone loss in FFI is highly localised to the mediodorsal and anterior ventral nuclei of the thalamus (Manetto *et al.*, 1992).

In addition to point mutations, numerous insertion mutations have been described which link to CJD. All these insertion mutations are found in the same region of the PrP gene, corresponding to between codons 51 and 91 of the human PrP gene. In unaffected individuals there are 5 repeated octapeptides in this region. In some individuals affected with CJD the number of these repeats is increased from 5 to 10, 11, 12, 13 and 14. All individuals with such insertions contract CJD but there is evidence that the number of repeats is related to age of disease onset and to the duration of disease (Goldfarb *et al.*, 1991; Owen *et al.*, 1991).

It may be that these mutations in the human PrP gene increase the probability of PrP spontaneously forming an infectious conformation, if such a conformation exists. This would be *de novo* generation of infectivity. Alternatively, it may be that these mutations predispose an individual to infection by an agent which is ubiquitous. In either case, the different PrP proteins produced by such mutations may influence the pathology and clinical outcome of disease.

**Table 2. Mutations associated with CJD**

| | |
|---|---|
| **Point** | 200 glutamine → lysine<br>178 asparagine → aspartic acid |
| **Octapeptide insertion (total number)**<br>Between codons 51-91 | 10<br>11<br>12<br>13<br>14 |

## 1.5 The PrP gene, its structure and expression

### 1.5.1 Conservation and structure

The PrP gene (also referred to as Prn-p in mice and PRNP in humans) has a structure of one or two 5' non coding exons which are separated from the open reading frame exon by a large intron (>10 kb). This arrangement is well conserved between species such as mouse (Westaway *et al.*, 1987, 1994c), hamster (Basler *et al.*, 1986), human (Puckett *et al.*, 1991) and sheep (Goldmann *et al.*, 1990). The PrP gene in these species appears to be single copy (Westaway and Prusiner, 1986; Oesch *et al.*, 1985; Liao *et al.*, 1986 and Goldmann *et al.*, 1990) and the chromosomal location for mice and humans has been mapped to chromosome 2 and to the homologous chromosome 20 respectively (Sparkes *et al.*, 1986).

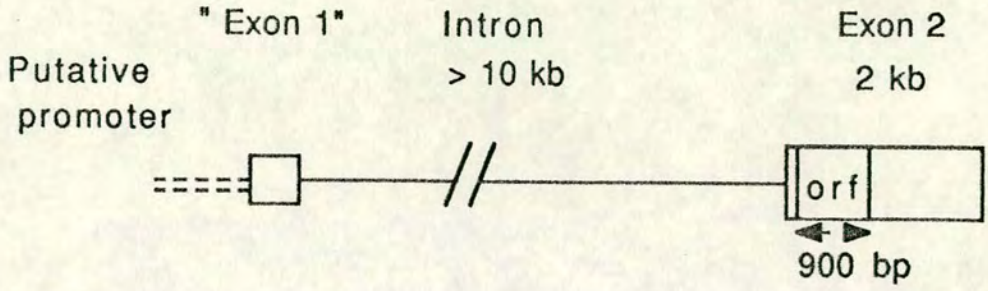



**Figure 1.**

Structure of rodent PrP gene as known in 1990. This figure is based on data by Basler *et al.*, (1986) and Westaway *et al.*, (1987). The basic gene structure is conserved between mouse and hamster. The entire open reading frame is contained within a single exon, dubbed exon 2. One 5' untranslated exon was shown to exist on the hamster PrP gene whereas at least one 5' untranslated exon was found on the mouse PrP gene. This exon was dubbed exon 1. The 5' exon was separated from exon 2 by a large intron, the size of which was at least 10 kb.

In addition to the structural similarities of the PrP gene between species there is also a high conservation at the nucleotide sequence level. Goldmann *et al.*, (1990) showed that the protein coding region of PrP mRNA is 80-90% homologous between sheep, mouse and hamster.

Based on Southern hybridisation Westaway and Prusiner (1986) have shown that sequences related to the hamster PrP gene may be present in lower organisms such as *Drosophila melanogaster, C.elegans and S.cerevisiae.*

However no homologous gene has been isolated in non-mammalian species.

### 1.5.2 5' non-coding exon (s) and promoter regions

As mentioned previously, at the start of this project only the gene structures for both hamster and mouse PrP genes were known. In the case of the mouse gene the 5' region of the gene was not well characterised. During the course of this project, however, it became clear that mouse, unlike hamster, had two 5' exons (Westaway *et al*, 1991, 1994 c). Later, the structure of the human PrP gene was elucidated and this too was found to have only one 5' exon (Puckett *et al.*, 1991). The significance of these differences and the function of the 5' exon (s) is not known.

The length of PrP exon one in humans is 136 bp and transcription is initiated at one site (Puckett et al., 1991). In hamsters transcription has been shown to be initiated at multiple sites which result in exon 1 having a variable length of 56-82 bp (Basler *et al.*, 1986).

Preliminary work by Westaway *et al* (1987) has shown that the murine PrP gene also has multiple transcriptional start sites. That the PrP gene can initiate transcription at multiple start sites is not unusual, indeed many genes of diverse function also do this: dihydrofolate reductase gene (Mitchell *et al.*, 1986), D1A dopamine receptor gene (Minowa *et al.*, 1992), growth hormone-releasing hormone gene (Gonzalez-Crespo and Boronat, 1991) and glycogen phosphorylase (Herrick *et al.*, 1993), to name but a few. The existence of multiple transcriptional start sites in these genes, and most others with multiple transcriptional start sites, can be attributable to their lack of a "TATA" box which has been shown to be involved in the accurate initiation of transcription (for review see Smale and Baltimore, 1989). Sequence analysis of the promoter regions of hamster (Basler *et al.*, 1986) and human (Puckett *et al.*, 1991) PrP genes has shown that they also do not contain a TATA box. Whether the use of differential start sites within the PrP gene serve any functional significance is not known, although there is evidence that alternative start sites in the alpha amylase gene serve a regulatory function. When expressed in the pancreas one site is utilised but when expressed in the salivary gland a different start site is used (Young *et al.*, 1981).

Our knowledge of the promoter regions of PrP genes in different species is limited. As yet, this knowledge is based almost entirely on sequence data. The preliminary work focused on the hamster gene (Basler *et al.*, 1986) and then the human PrP gene (Puckett *et al.*, 1991) . This year, however, sequence data from the 5' flanking region of sheep (Westaway *et al.*, 1994 b) and mouse PrP genes has been published (Westaway *et al.*, 1994 c). In human, only 71 bp of the 5' flanking region has been sequenced and no functional analysis of this region was carried out. Despite this, the sequence data has shown that this region has no "TATA" box, is very GC rich (83%) and has several potential binding sites for the transcription factor Sp1. Larger regions of hamster, mouse (this study) and sheep 5' flanking regions have been sequenced and by comparison with the human PrP promoter we can see that these features are also conserved in these regions. In addition to the sequence data, Basler *et al* (1986) showed that a subclone which contained approximately 5 kb of hamster PrP 5' flanking region can function as a promoter. This was done in an *in vitro* transcription assay and by comparison, the adenovirus late promoter was shown to be approximately 10 fold stronger.

### 1.5.3 Introns: possible functions

From recently published data (Westaway *et al.*, 1994 c) it is now known that the murine PrP gene has three exons which are separated from each other by two introns. While the size of intron 1 has been shown to be approximately the same in different strains of mice, the size of intron 2 is dependent on the strain of mouse (Westaway *et al.*, 1994 c), specifically on the PrP allele (either $PrP^a$ or $PrP^b$) . These alleles differ from each other by two amino acids at codons 108 and 189 (Westaway *et al.*, 1987) and these differences are consistently found in mice of different (see section 1.6) *Sinc* genotypes (s7 or p7). Mice which harbour the $PrP^a$ allele (*Sinc* s7) have been found to have an intron 2 size of approximately 17.3 kb whereas mice which

harbour the $PrP^b$ allele (*Sinc* p7) have a deletion in this intron of about 6.7 kb (Westaway *et al.*, 1994 c).

The effect of intron sequences in the regulation of other genes is well documented. In some cases introns were shown to have a positive or stimulatory effect on gene expression such as those found in the dihydrofolate reductase gene (Schmidt *et al.*, 1990) and thymidine kinase gene (Rotheneder *et al.*, 1991). In contrast, introns have also been shown to have a negative or inhibitory effect on gene expression e.g introns in the c-fos gene (Lamb *et al.*, 1990) and in the alpha 1(I) collagen gene (Bornstein and McKay, 1988).

Whether the intron (s) of the PrP gene have a role in the control of gene expression is not known as it has not been adequately addressed. There is , however, one study (Scott *et al.*, 1989) which provides evidence, albeit indirect, that the intron of the hamster PrP gene may have an important effect on PrP gene expression. In this study, transgenic mice were generated using an intronless minigene construct. The minigene contained 1.5 kb of hamster PrP promoter, the open reading frame and the polyadenylation signal for the PrP gene. The mice which were transgenic for this construct failed to express hamster PrP protein in their brains. Mice which were transgenic for a genomic cosmid clone of the hamster PrP gene did produce both hamster PrP mRNA and protein in their brains. One interpretation of this data may be that the intron of the hamster PrP gene is crucial in directing high level gene expression. This could be achieved by the presence of a strong enhancer in the intron. It is impossible, however, to say with a degree of certainty that the minigene construct failed to express in mouse brain because it lacked intron sequences. There may be other explanations for this, such as transgene integration into a transcriptionally repressed site or absence of correct 5' flanking sequences. However, an investigation into the role of intron sequences on PrP gene expression would seem worthwhile.

### 1.5.4 The open reading frame exon: codes for PrP protein and 3' UTR

In all species yet examined the entire open reading frame of the PrP gene resides on one exon and is not interrupted by introns (Oesch *et al.*, 1985; Westaway *et al.*, 1987; Puckett *et al,* 1991). In hamster and mouse the ORF exon is approximately 2000 bp in length (Basler *et al.*, 1986; Westaway *et al.*, 1987) whereas sheep have a much larger ORF exon of approximately 4000 bp (Goldmann *et al.*, 1990). The actual open reading frame of PrP between these species are very similar in size (-760 bp), but the remainder of the exon is composed of 3' untranslated sequence. There is a 10 bp untranslated region between the 5' splice site and the translational initiation codon ATG and this is well conserved between mouse (Westaway *et al.*, 1987), hamster (Basler *et al.*, 1986) and sheep (Goldmann *et al.*, 1990). An interesting feature about the ORF exon is that the 5' end is organised into repeats. In rodents there are two small repeats (18 bp) which are followed by five 27 bp repeats. In humans, variation in the number of these repeats has been shown to correlate strongly with some cases of familial CJD (Goldfarb *et al.*, 1991; Owen *et al.*, 1991)
See figures 2, 3 and 4 for comparisons of PrP open reading frames of mouse, hamster and human. From such comparisons, the presence of the repeated regions can clearly be seen (see figure legend for details)

The 3' untranslated region of the ORF exon is not well characterised, but from sequence data one can see that it contains some highly conserved areas between hamster, mouse, sheep and human (Goldmann *et al.*, 1990). The 3' UTR may contain some regulatory regions involved in, for example, message stability or RNA processing. An example of this is found in the 3' UTR of mouse protamine 2 mRNA. An 18 kDa protein when bound specifically to the 3' UTR is known to cause translational repression (Kwomdel and Hecht, 1993).

**Figure 2.**

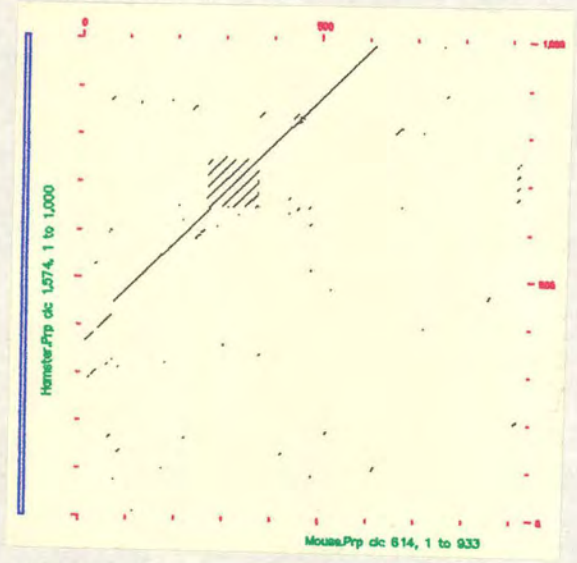



**Figure 3.**

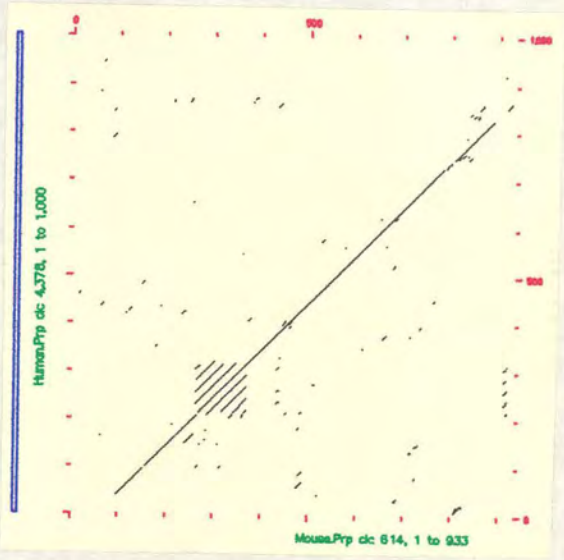
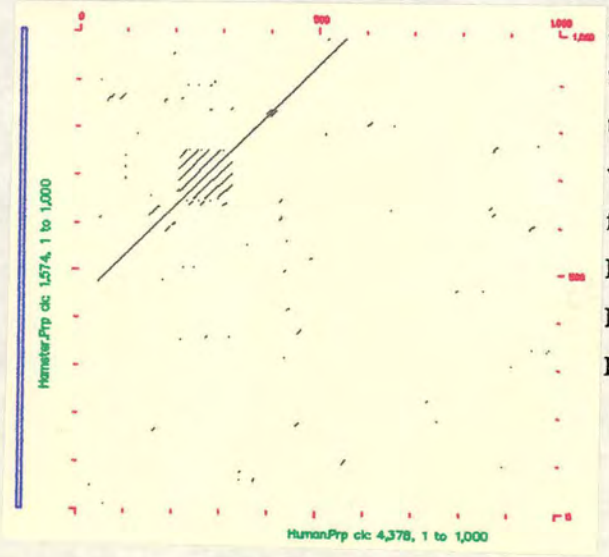



**Figure 4.**




These figures represent graphical comparisons between two sequences and are generated by a programme called DOTPLOT (Maizel and Lenk, 1981). This programme constructs dot plots at points of homology between two sequences. In the case of mouse, hamster and human PrP open reading frames this homology is so high that these individual dots merge to give an almost unbroken straight line. Repeated regions can also be visualised using DOTPLOT and these can be seen in each figure as a series of short, unbroken lines which form a characteristic square shape around the central line.

Figure 2 represents the comparison between mouse (Westaway *et al.* ,1987) and hamster ORFs.

Figure 3 represents the comparison between mouse and human (Puckett *et al.*, 1991) ORFs.

Figure 4 represents the comparison between human and hamster (Basler *et al.*, 1986) ORFs.

### 1.5.5 Expression of PrP mRNA in different tissues

There have been a number of studies, using different detection methods, which have determined in which tissues PrP is transcribed. The three main detection methods used were RNA slot blots, Northern blot and *in situ* hybridisation.

Initial work centred on the detection of hamster mRNA in a variety of tissues using Northern hybridisation and RNA slot blots. Using Northern hybridisation, Oesch *et al.*, (1985) showed that single species of PrP transcripts are found in the brain, heart and lung and that they all appear to be the same size. Brain expressed PrP more highly than the other two tissues. By using the slot blot procedure PrP transcripts were found in brain, pancreas, heart, lung, kidney, testes and to a lesser extent in spleen. Caughey *et al.*, (1988) also found PrP transcripts in a variety of tissues and cell lines and from Northern blots estimated that spleen contained about 0.8 % the amount of PrP transcripts when compared to brain. This result was confirmed by Brown *et al.*, (1990) using *in situ* hybridisation. This revealed that PrP was barely detectable in liver and undetectable in spleen. That mRNA is barely detectable in spleen is of interest as spleen has been shown to harbour a high titre of scrapie agent in infected animals (Eklund *et al.*, 1967). By using antisense RNA probes for *in situ* hybridisation, Manson *et al.*, (1992 b) detected PrP mRNA in the majority, perhaps all, neurones of the brain. Different types of neurones, however, were shown to have different levels of PrP within them. In the same study PrP protein, as detected by immunocytochemistry, was found to be expressed in some, but not all, neurones of the brain. This therefore suggests that regulation of PrP gene expression may be controlled, at least in part, at a post-transcriptional level.

In all these studies there has been no detectable difference in the amount of PrP mRNA between diseased and undiseased animals. This is in contrast to the levels of PrP protein which have been reported to rise between 10-100 fold during the disease course (Bruce *et al.*, 1989).

Table 3. Differential expression of PrP mRNA.

| | |
|---|---|
| **Brain** | + + + |
| **Heart**<br>**Lung**<br>**Pancreas** | + + |
| **Spleen**<br>**Liver** | + |

+ + + denotes strong expression

+ denotes weak expression

### 1.5.6 Developmental expression of PrP mRNA

McKinley *et al.*, (1987) were the first to examine whether PrP gene expression is developmentally regulated, in hamsters. They found using Northern blotting, that PrP transcripts were only detectable post-natally, rising from day 1 to day 20. At day 20 these levels became stabilised and were maintained throughout adult life. In rats PrP mRNA could be detected just before birth (embryonic day 20). These transcripts increased 3 fold to adult day 70 and this developmental change was shown to be due to a 3 fold increase in the transcription rate as measured by nuclear run off assays (Lieberburg , 1987).

In mouse, Lazarini *et al.*, (1991) showed that PrP was developmentally regulated at a post-natal level. PrP mRNA could be detected at birth, by Northern blot, and then increased 4 fold to day 20 where it was maintained at this level throughout adult life. In one study, Manson *et al.*, (1992 a) showed that PrP mRNA can be detected prenatally in mice. *In situ* hybridisation analysis detected PrP transcripts in the developing embryonic

brain and spinal cord by 13.5 days but none were detected at day 6.5 or day 9.5. In addition PrP transcripts were also found in the peripheral nervous system by day 16.5 and also in specific non-neuronal populations.

In addition to developmental regulation, and the regulation which controls the level of PrP gene expression in different tissues, PrP has also been shown to be regulated by the action of nerve growth factor (NGF). In rat neuroblastoma cell line PC12, NGF was shown to increase the levels of PrP mRNA. This increase was detectable by 2 days but maximal increase was observed after 7 days (Wion *et al.*, 1988). Interestingly beta amyloid precursor protein (APP) mRNA, which like PrP forms amyloid plaques during disease (Masters *et al.*, 1985), was also shown to be affected by the action of NGF, although not in the same way. Mobley *et al.*, (1988) provided *in vivo* evidence that injection of NGF can regulate both PrP and APP mRNA in neonatal hamsters.
The DNA sequences and transcription factors involved in the regulation of PrP gene expression are unknown. However, it appears likely that PrP can be regulated at a transcriptional, post-transcriptional and possibly (during disease) at a post-translational level.

## 1.6 Transgenetic experiments

The development and analysis of transgenic mice has proved invaluable for studying many facets of the transmissible degenerative encephalopathies. These facets include the species barrier, relationship of PrP to the causative agent, the effect of mutations in the PrP gene and in the control of scrapie incubation period. A summary of several of these transgenic experiments are described below.

Hamster PrP transgenic mice

In 1989 the first PrP transgenic mouse lines were developed. The mice harboured copies of the Syrian hamster PrP gene and were designated as being tg (Ha PrP) mice. Three transgenic lines were produced and contained varying copy numbers of the hamster transgene. Tg 69 and 71 had 4-8 copies of the hamster PrP gene while tg 81 harboured 30-50 copies. On inoculation with hamster specific scrapie agent (sc237) the tg (Ha PrP) mice all developed disease. Tg 81 was found to have the shortest incubation period - ~ 75 days, whereas the tg lines with less transgene copies had scrapie incubation periods of ~ 170 days. Non-transgenic mice did not develop disease 500 days after inoculation. The tg (Ha PrP) mice had $PrP^{sc}$ present in their brains, as determined by proteinase k resistance. In addition, the pathology was characteristic of hamster scrapie (Scott *et al.*, 1989). These results suggested that the PrP gene can modulate scrapie susceptibility (species barrier), incubation times and neuropathology. In addition, this experiment showed that the transgene copy number also can play an important role in the outcome of disease.

GSS 102 transgenic mice

As mentioned in a previous section, some patients who are affected by GSS have a mutation at codon 102 of the PrP gene which causes a proline to lysine change. In an attempt to develop an experimental model of GSS, transgenic mice were generated by Hsiao and co-workers (1990) which harboured an analogous mutation (mouse codon 101) on the PrP gene. The transgene copy numbers varied from approximately 10 - 64 copies. These mice were found to develop spontaneous neurodegeneration at an average age of 166 days. They presented with clinical signs such as ataxia, lethargy and rigidity and gradually deteriorated until death. At a pathological level, vacuolar degeneration and astrocytic gliosis were seen. However, no amyloid plaques were detected. $PrP^{sc}$ was present in the brains of the transgenic mice but at very low levels.

One of the main problems in the interpretation of this experiment is the effect of high transgene copy numbers. For example, this experiment does not properly address whether spontaneous disease is caused by the 101 mutation or by increased PrP copy numbers. The transgenic experiments of Scott *et al.*, (1989) clearly indicate that the transgene copy number can influence certain aspects of disease (see above). It seems likely that this contributed to the pathogenesis of spontaneous neurologic disease, as reported by Hsiao *et al.*, (1990) as the mean age of disease onset correlated inversely with transgene copy number.

Although technically more difficult, generation of transgenic mice by homologous recombination would act as a control for copy number effects as only one copy of a gene would be altered at a given time.

During the course of this study two developments heightened the need to understand the regulation of the PrP gene. These developments were a) the production of PrP overexpressing mice which developed neurodegeneration and b) the production of PrP null mice, which were subsequently shown to be resistant to scrapie.

These studies are discussed below.

Overexpressing PrP mice

A recent study by Westaway *et al.*, (1994 a) has provided evidence that high copy numbers of wild type (wt) PrP transgenes are pathogenic, resulting in degeneration of skeletal muscle, peripheral nerves and CNS.

A number of different transgenic lines were used in this study and differed from each other in terms of transgene copy number, transgenic status (i.e hemizygous or homozygous), and by the PrP transgene which they carried. Some mice were transgenic for the sheep PrP gene (e.g tg[sheep $PrP^{o/+}$ ] ), the Syrian hamster PrP gene (e.g tg[sHa $PrP^{+/+}$ ] ) and also for different alleles of the mouse PrP gene (PrP-B) which confer long

incubation periods on mice with most, but not all, strains of scrapie (e.g tg[mo PrP-$B^{0/+}$ ] ) as described by Westaway *et al.*, (1987, 1991). The transgene copy number of these mice was estimated to vary from 3-120.

The animals with 31 copies, or more, of the PrP transgene all developed, spontaneous neurologic disease. The clinical symptoms included tremor, kyphosis, gait abnormalities and in tg(mo PrP-$B^{0/+}$ ) mice, hind leg paralysis. Onset of disease usually occurred in older mice (~ 560 day in tg[mo PrP-$B^{0/+}$ ] mice and > 600 days in tg[sheep $PrP^{0/+}$ ] ). However, onset of disease was detected as early as 220 days in the 120 copy number sHa PrP homozygous mouse.

All of these mice had vacuolation in the CNS and presented with necrolizing myopathy and peripheral neuropathy. Myopathy was most striking in the quadriceps, although it was found in all types of skeletal muscle including intercostal muscles and diaphragm. The vacuolation was localised to the stratum lacunosum, molecular layer of the hippocampus, midbrain tegmentum and to the superior colliculus.

In the tg (mo PrP-$B^{0/+}$) line, which contained 31 transgene copies, PrP mRNA was found to be overexpressed by ~ 7 fold , while $PrP^{c}$ was overexpressed by ~ 8 fold in the brain. No protease resistant PrP was found in the mice, in either the brain or quadriceps muscle. This study showed that elevated PrP expression, caused by transgene dosage, is pathogenic and produces a progressive neurologic disease.

PrP null mice

To address the normal function of PrP and the role it plays in the transmissible degenerative encephalopathies PrP null mice have been generated. PrP homozygous null mice ($PrP^{0/0}$ ) have been created in two labs, using two different targeting strategies (Bueler *et al.*, 1992; Manson *et al.*, 1994). Although both strategies resulted in mice expressing no PrP protein, Bueler and co-workers found that their null mice produced a neomycinphosphotransferase/PrP transcript. The null mice of Manson and co-workers did not produce any PrP related mRNA sequences. Whether this difference will manifest itself at a phenotypic or disease related level is not known at this stage.

Both laboratories have, however, found that both heterozygote and homozygote null mice show no obvious phenotype. They show normal development up to the age of 7 months and are fertile. Apart from the very recently described synaptic differences, described in section 1.2 (Collinge *et al.*, 1994), the null mice are apparently normal.
That a more obvious phenotype is not apparent is perhaps surprising considering that the PrP gene is expressed in different embryonic (Manson *et al.*, 1992) and adult tissues (Oesch *et al.*, 1985) and is highly conserved in all mammals as yet studied (Goldmann *et al.*, 1990). The generation of knock-out mice for other genes producing no obvious abnormality is not without precedent, however. Ablation of the p53 gene resulted in mice which showed normal development and reproduction, when a lethal phenotype was expected (Donehower *et al.*, 1992).

Preliminary experiments have been conducted in which PrP null mice (both heterozygotes and homozygotes) have been challenged with scrapie agent. $PrP^{0/0}$ have been shown to be completely protected against scrapie for at least 2 years while all wild type controls ($PrP^{+/+}$) died within 6 months. Heterozygotes ($PrP^{o/+}$) were shown to be partially protected from scrapie; having incubation periods of around 10 months (Bueler *et al.*, 1993). This study reinforces the idea that the levels of $PrP^c$ and the incubation period are inversely correlated. It also strongly suggests that the presence of PrP is necessary for disease.

## 1.7 Aims of this thesis

The aims of the work described in this thesis are as follows:

1) To construct a cosmid genomic library using mouse DNA. From this library a PrP clone will be selected and analysed in detail.

2) Analysis of a PrP cosmid clone would consist of determining a simple restriction map and then identifying the 5' untranslated exons and the 5' flanking region.

3) This 5' region of the PrP gene would be nucleotide sequenced and its homology to other published sequences determined. In addition, this sequence would be scanned for potential transcription factor binding sites. These may indicate the presence of regulatory elements.

4) Having identified the 5' flanking region through sequence homology, the next stage would be to analyse the promoter by linking it to a reporter gene and transfecting into a suitable cell line. Deletions will then be made in the promoter region with the aim of identifying regulatory elements which are important for the expression of the PrP gene.

5) Finally, the transcriptional start sites, of the PrP gene, will be determined by primer extension and mapped onto the promoter.

## 2.1 Chemicals and enzymes

All chemicals and enzymes were purchased from Sigma, Gibco-BRL, Boeringer-Mannheim, Promega and Pharmacia and stored according to the manufacturers instructions. Packaging mixes and cosmid vector were purchased from Stratagene. Distilled water ($dH_2O$) was used routinely. However, for tissue culture and RNA analysis water was purified from a Millipore milli-QUF system.

## 2.2 Strains of E.coli used in this study.

**Table 4.**

| Strain | Genotype |
|---|---|
| NM554 | *recA13, araD139, Δ(ara-leu)7696 Δ(lac) 17A, gal a, gal k, hsdr, rpsl(strr), mcr A, mcr B* |
| CMK 603 | *thr, leu, thi, supE, recBC, T1R, T5R, R-, M+, lacZΔM15, lacY, F'lacq1, lacZM15 pro+* |
| DH5α | *F'/endA1, hsdR17(rk-mk+), supE44, thi-1, recA1, gyrA(Nalr), relA1, Δ(lacZYA-argF)U169, (ϕ80dlacΔ(lacZ)M15* |

Strain NM554 was used for the cosmid cloning, whereas strain CMK 603 and DH5α were used for routine transformations.

The recipes for the solutions found in this chapter and in subsequent chapters can be located in the appendices. Most of these solutions were autoclaved before use to aid in the prevention of nuclease contamination and were stored in small batches.

## 2.3 Growth media and conditions used for culturing bacteria

Bacteria were plated on L-agar and incubated upside down, at 37°C, overnight. For liquid culture a single bacterial colony was inoculated into the appropriate volume of L-broth, usually 3-5 ml and incubated with shaking at 37°C for 6 - 16 hrs. When appropriate, antibiotics ampicillin and kanamycin were added to the medium at 100 $\mu g\ ml^{-1}$ and 50 $\mu g\ ml^{-1}$ respectively.

## 2.4 Techniques of DNA manipulation

### 2.4.1 Small scale preparations of plasmid DNA

These small scale preparations of plasmid DNA are more commonly known as mini-preps. Two different types of mini-preps were used. Both are based around the original method of Birmboim and Doly (1979). The first one is known as the alkaline lysis method and was generally used for screening transformants. The second method (magic mini-prep - Promega) has an additional step which ensures much purer DNA and this method was generally used for preparing template for use in double stranded DNA sequencing.

i) Alkaline lysis

a) 3 ml of L-broth (plus antibiotic) was inoculated with a single bacterial colony and incubated at 37°C, with shaking, for at least 6 hrs but usually overnight.

b) 1.5 ml of the culture was spun down in a microfuge for 30 sec at 13000 rpm and the supernatant discarded.

c) The pellet was vortexed in 100 µl of solution I.

d) To this was added 200 µl of freshly made solution II.

e) 150 µl of solution III was then added and mixed by vigourous vortexing. After being stored on ice for 5 mins the solution was spun for 5 mins at 13000 rpm in a microfuge and the DNA containing supernatant was removed into a fresh tube.

f) The DNA was ethanol precipitated and resuspended in 50 µl of T.E.

ii) Magic mini-prep (Promega)

a) A bacterial pellet was prepared as above.

b) The pellet was resuspended by vortexing with 200 µl of cell resuspension solution.

c) To this was added 200 µl of cell lysis solution (identical to solution II) and this was mixed by inverting the tube several times.

d) After lysing the bacteria, the solution was neutralised by adding 200 µl of neutralisation solution. This was mixed by inversion and spun down in a microfuge at 13000 rpm for 5 mins.

e) 1 ml of magic mini-prep purification resin was added to the supernatant and gently mixed.

f) The DNA / resin mix was then pushed through a magic mini-column. This was prepared by removing the plunger from a disposable 2 ml syringe and attaching a luer-lock extension of the mini-column to the syringe barrel. After the DNA had been pushed through, the column was removed and the plunger withdrawn. The syringe barrel was then re-attached to the mini-column.

g) 2 ml of wash solution was pipetted into the barrel and pushed through. The mini-column was then dried by a brief spin in a microfuge at 13000 rpm.

h) 50 µl of $H_20$ or T.E was then added to the mini-column and spun for 1 min at 13000 rpm. The purified DNA was collected into an Eppendorf.

### 2.4.2 Large scale preparations of plasmid DNA

Two methods were used to isolate and purify plasmid DNA in large amounts. The first method was based on the Birmboim and Doly (1979) alkaline lysis method and is essentially a scaled up mini-prep. This method was used for most plasmid preps. The second method employs SDS and was specifically used to isolate cosmid clones, as it was more gentle. This is a standard method for isolating large plasmids (Sambrook *et al.*, 1989).

Both methods incorporate the use of cesium chloride gradient centrifugation and the difference lies in the methods by which the bacteria are lysed.

a) Bacterial lysis

i) Alkaline lysis method

1) A bacterial pellet was obtained from a 500 ml overnight culture by centrifugation at 7000 x g for 15 mins at 4°C .

2) This pellet was resuspended in 18 ml of ice-cold solution I. 2 ml of a 10 mg $ml^{-1}$ lysozyme solution (in 10 mM Tris-HCl pH 8) was added and gently mixed.

3) To this was added 40 ml of fresh solution II in order to lyse the cells.

4) 20 ml of ice-cold solution III was then added and mixed well. This was left on ice for at least 10 mins. This was centrifuged at 7000 x g for 15 mins at 4°C.

5) The supernatant was filtered through a layer of cheese cloth to remove any residual bacterial debris and mixed with 0.6 vol isopropanol. This was left for 10 mins. The nucleic acids were recovered by centrifugation at 7000 x g for 15 mins at room temperature.

6) The pellet was allowed to dry for about 15 mins.

ii) SDS method

1) The bacterial pellet from a 500 ml overnight culture was resuspended in 10 ml of an ice-cold solution of 10 % sucrose in 50 mM Tris-HCl pH 8.

2) 2 ml of a 10 mg $ml^{-1}$ lysozyme solution in 10 mM Tris-HCl pH 8 was added and mixed and then transferred to a 30 ml screw-cap centrifuge tube.

3) 4 ml of 0.5 M EDTA pH 8 was mixed into the suspension and placed on ice for 10 mins. 4 ml of 10 % SDS solution was then added and mixed immediately by gentle inversion.

4) 6 ml of 5 M NaCl was then added and again gently mixed. This was stored on ice for 1 hr.

5) This was then centrifuged at 65000 x g for 30 mins at 4°C and the supernatant transferred to a fresh tube where it was phenol:chloroform extracted.

6) The aqueous phase was then ethanol precipitated and the DNA recovered by centrifugation at 7000 x g for 20 mins at 4°C.

### 2.4.2 b) CsCl ethidium bromide gradient centrifugation

The nucleic acid pellet was dissolved in CsCl (13 ml; 1g $ml^{-1}$) solution. 1 ml of a 10 mg $ml^{-1}$ ethidium bromide solution was added and mixed well. The ethidium bromide/protein scum which forms was removed by centrifugation at 5000 x g for 5 mins at room temperature. The clear red supernatant was then transferred to a Beckman "quick seal" tube and sealed.

The sealed tubes were then centrifuged at 350 000 x g for 16 hrs at 20°C in a NVT65 near vertical rotor or at 100 000 x g for 48 hrs in a 70.1 Ti rotor, also at 20°C.

After centrifugation two bands were present in the middle of the tube. The lower band containing the supercoiled plasmid DNA was removed, by side puncture, using a 19 gauge needle attached to a 2 ml syringe. Another needle was inserted into the top of the tube to equalise the pressure.

### 2.4.2 c) Rebanding plasmid DNA

Plasmid DNA was subjected to a second round of centrifugation when it was to be used for transfections into mammalian tissue culture cells. This removed any residual chromosomal DNA or RNA and any nicked or linear plasmid which may affect transfection efficiency. This process involved removing the plasmid band and transferring it directly to a fresh tube. The tube was then topped up with the CsCl solution (1 g $ml^{-1}$) and spun again in exactly the same fashion as previously described.

### 2.4.2 d) Removal of ethidium bromide from purified plasmid DNA

After ultra-centrifugation, the ethidium bromide / CsCl was removed from the plasmid DNA solution by adding an equal volume of T.E saturated butan-2-ol. After inverting the closed tube several times the two phases were allowed to separate. The ethidium bromide partitioned into the upper butan-2-ol phase and was discarded. The process was repeated until the bottom phase lost its colour.

### 2.4.2 e) Removal of CsCl from purified plasmid DNA

CsCl was removed from the DNA solution by dialysis against 5 x 1L of either T.E (pH 8) or MilliQ $H_2O$. This was carried out at $4^0$C over a period of 24 hrs.

### 2.4.2 f) Spectrophotometric determination of DNA concentration

The optical density at 260nm ($OD_{260nm}$) was measured in a quartz cuvette in a spectrophotometer. An OD of 1 corresponds to 50 $\mu g\ ml^{-1}$. The length of the light path was 40 mm.

### 2.4.3 Digestion and agarose gel electrophoresis of DNA

DNA was digested using 1 unit of enzyme per microgram of DNA in the presence of the appropriate restriction enzyme buffer. When necessary the total volume was made up with $dH_2O$. The reaction was incubated at 37°C, except for enzymes whose optimal temperature was not 37°C e.g SmaI. In both cases, however, the incubation was for 1 - 3 hrs.

0.3 - 1.5 % agarose gels were used, depending on the size of the DNA fragments which were to be separated. In all cases 1 x TAE solution was used as the running buffer and was also used in the composition of the gel.

1 µl of gel loading buffer was added to, usually, 10µl DNA before being loaded into the gel. 1 µg of HindIII digested Lambda DNA was used as size standards. Gels were electrophoresed at 15-100 V for between 1-16 hrs depending on the experiment. After electrophoresis the gel was soaked in 400 ml of a 0.5 µg $ml^{-1}$ solution of ethidium bromide for 10-60 mins. The ethidium bromide was then decanted and replaced with 400 ml of $dH_2O$. This was washed at room temperature with shaking for 15 mins to destain the gel. The DNA was viewed using a U.V transilluminator and photographed using Polaroid 667 film.

### 2.4.4 Estimation of restriction fragment size

HindIII digested Lambda DNA were used as size standards in every gel.

The migration distance of the λ DNA markers were measured and plotted against the $log_{10}$ of their molecular weight. From this standard curve the migration distance of a restriction fragment was converted to its apparent molecular weight.

### 2.4.5 Isolation and recovery of restriction fragments from agarose gels

Two techniques were used for the recovery of restriction fragments from agarose gels:

a) Centrifugation through blotting paper.

b) Low melting point agarose gels.

**Centrifugation through chromatography paper**

This technique involves cutting out a band visualised in an ethidium bromide stained agarose gel and centrifuging this through blotting paper (Weichenham, 1991). Two 1.5 ml Eppendorfs had their lids removed and one was placed inside the other. The bottom of the upper tube was punctured with using a 19 gauge needle. A small square (2 cm x 2 cm) of chromatography paper e.g Whatman 3MM, was wrapped around the inside of the upper tube to form a tight fitting cup. Into this cup was placed the gel piece containing the restriction fragment of interest. The assembly was then centrifuged at 13000 x g for 1 min using an adaptor fitted into the JA-20 rotor. After centrifugation the DNA was forced through the blotting paper into the lower tube, along with the electrophoresis buffer. The dehydrated agarose remains in the blotting paper. The DNA was then ethanol precipitated and resuspended in a suitable volume of T.E or $H_20$. This was the method most commonly employed to isolate restriction fragments.

**Low melting point agarose**

Low melting point (LMP) agarose gels were poured and run under the same conditions as ordinary gels except that the voltage was always low (25 V) to prevent the gel melting. The band of interest was cut out using a sterile scalpel blade and placed in a 15 ml tube. 5 vols of 20 mM Tris-HCl (pH 8), 1 mM EDTA (pH 8) were mixed per 1 vol of agarose and incubated at 65$^0$C until the agarose melted. The mixture was then phenol extracted, phenol:chloroform extracted and then finally chloroform extracted. The phases were separated between each extraction by centrifugation at 1000 x g for 10 mins at room temperature. The aqueous phase was ethanol precipitated and the DNA recovered by centrifugation at 10 000 x g for 30 mins at 4$^0$C.

### 2.4.6 Phenol:chloroform extraction of nucleic acids

Proteins were removed from DNA solutions by carrying out a phenol:chloroform extraction. An equal volume of phenol was added to the nucleic acids solution and vortexed for 10 secs. This was then spun at top speed in a microcentrifuge for two mins to separate the phases. The aqueous phase was then transferred to a fresh tube and an equal volume of chloroform was added and vortexed for 10 secs. Again the solution was centrifuged at top speed in a microcentrifuge and the nucleic acid containing aqueous phase was removed for ethanol precipitation.

### 2.4.7 Ethanol precipitation of nucleic acids

2-4 vols of absolute ethanol and 0.1 vol of 3 M sodium acetate was added to the DNA. This was vortexed and stored at -20$^0$C for 30 mins. The DNA was recovered by centrifugation at 13 000 rpm in a microfuge for 10 mins, followed by washing the pellet in 70% ethanol.

### 2.4.8 DNA modifications

#### a) Filling in 3' termini

Up to 5 µg of restriction enzyme digested DNA was used. 1 µl of a fresh 1 mM solution of all four dNTPs (dATP, dGTP, dCTP, dTTP) was added along with 1 µl (6 units) of the Klenow fragment of DNA polymerase I. This was vortexed briefly and incubated at room temperature for 30 mins. To inactivate the enzyme(s) the tube was heated at 75$^0$C for 10 mins.

**b) Removal of protruding 3' termini**

Klenow fragment of DNA polymerase I has a 3'→5' exonuclease activity which can be used to remove protruding 3' termini, but T4 DNA polymerase is much more efficient at this process and hence was used. Again up to 5 μg of DNA was restriction digested and to this was added 1 μl of a 2 mM dNTP solution. 1 unit of T4 DNA polymerase was added per μg of DNA, mixed and incubated for 15 mins at $12^0$ C. The enzymes were inactivated by heating at $75^0$ C for 10 mins.

**c) Dephosphorylation of 5' phosphate groups**

To prevent self ligation of vector molecules during cloning, the 5' phosphate groups were usually removed from linear DNA before sticky ended ligation. Calf intestinal alkaline phosphatase (CIP) was the enzyme used for all dephosphorylation reactions. The appropriate amount of DNA was restriction digested then phenol:chloroform extracted and ethanol precipitated. The recovered DNA was dissolved in 90 μl 10 mM Tris-Cl (pH 8.3). 10 μl of the 10 x dephosphorylation buffer was then added. For protruding termini 1 μl of CIP was added and the reaction was incubated for 30 mins at $37^0$C. For recessed termini 1 μl of enzyme was added and incubated at $37^0$C for 30 mins followed by 45 mins at $55^0$C with another addition of enzyme. The CIP was then inactivated by heating at $75^0$C for 10 mins and removed from the reaction by a phenol:chloroform extraction and ethanol precipitation.

### 2.4.9 Ligation reactions

#### a) Sticky-end ligations

Restriction fragments were gel purified, see section 2.4.5, and the amount of DNA recovered was estimated by comparison of band intensity with λ Hind III markers when run on the same ethidium bromide stained gel. At least a three fold molar excess of insert was ligated to the CIP treated vector in the presence of 1 µl T4 DNA ligase (1 unit) and 2 µl of 5 x ligation buffer. The DNA concentration of the ligation reaction varied but was usually around 1µg $ml^{-1}$. The volume was made to 10 µl with $dH_2O$ and mixed. This was then incubated overnight at 16°C.

#### b) Blunt-end ligations

Fragments were purified and quantified as described in sections 2.4.5 and 2.4.9 a. A higher insert to vector molar ratio was used in blunt-end ligations (>15 fold). 1 µl of high concentration T4 DNA ligase (6 units), 2 µl of 5 x T4 ligation buffer and $H_20$ to 10 µl were added, mixed and incubated overnight at 16°C. The ligation reactions were then transformed into E.coli.

### 2.4.10 Transformation of E.coli with plasmid DNA

1) Various strains of E.coli were used for the transformation procedure (see table 4). E.coli strains CMK and DH5α were regularly used but NM554 was also used. Transformations were carried out using the calcium chloride method of Cohen *et al.*, (1972). A single fresh colony was used to inoculate 100 ml of L-broth and this was incubated at 37°C with vigourous shaking until the cells were growing in approximately mid-exponential phase. This was estimated by determining the OD at 600 nm. Optimal $OD_{600nm}$ for CMK was empirically determined to be 0.2 - 0.3

whereas the optimal OD for NM554 and DH5α was 0.4 - 0.6.

2) Once the optimal OD was reached the culture was split into two 50 ml tubes and spun at 1000 x g for 10 mins at 4°C. These pellets were resuspended in 10 ml of 0.1 M ice-cold $CaCl_2$ and spun again for 10 mins. This time the pellets were resuspended in 2 ml ice-cold 0.1 M $CaCl_2$ and stored on ice. These competent bacteria were either used immediately or stored on ice overnight at 4°C.

3) Between 0.1 - 50 ng of DNA or ligation mix was mixed with 200 µl of competent cells and stored on ice for 30 mins to allow the DNA to precipitate onto the cell surface. By heat-shocking for 90 sec at 42°C and then storing on ice again the DNA can enter the bacterial cell. 800 µl of L-broth was added and the cells were shaken gently for 45 - 60 mins at 37°C. This allows the bacteria to express the antibiotic resistance gene before selection on ampicillin agar plates. After incubation at 37°C the culture was transferred to Eppendorf tubes and the bacteria recovered by centrifugation at 13000 rpm in a microfuge for 30 secs. 800 µl of supernatant was removed and the pellet was resuspended in the remaining 200 µl and plated out onto a L-agar plate containing the appropriate antibiotic. Plates were incubated overnight at 37°C.

Transformation efficiencies could be calculated from the number of transformants derived from using a given amount of DNA. For example, 0.01 ng and 0.1 ng would be in a transformation and the number of colonies present on the plate the next day could be counted. Transformation efficiencies are usually expressed in number of transformants per microgram of DNA. With this method, transformation efficiencies exceeded $1x10^6$ transformants $\mu g^{-1}$ of DNA.

### 2.4.11 Radioactively labelling DNA with $^{32}P$

#### a) Random primer extension

This method was used for labelling restriction fragments of double stranded DNA for use as probes (Feinberg and Vogelstein, 1983).

#### i) Preparation of DNA

DNA was restriction digested and run on a low melting point agarose gel. The desired band was cut out and the gel slice weighed in a microfuge tube. To it was added 3 ml of $H_20$ per gram of agarose and this was boiled for 10 mins. The DNA / agarose was aliquoted out into 30 - 50 ng amounts and stored at -20°C until needed.

#### ii) Preparation of oligonucleotide labelling buffer (OLB)

This buffer was a mixture of three separate solutions : A, B and C which were mixed together at a ratio of 100:200:150 respectively.
Once made OLB was stored in 10 µl aliquots at -20°C.

#### iii) Labelling reaction

30 - 50 ng of the DNA to be labelled was denatured by boiling for 10 mins then placed at 37°C for a few mins. To it was added 10 µl OLB, 1µl Klenow fragment,

5 µl (50 µCi) [$\alpha$-$^{32}$P] dCTP at 3000 Ci mmol$^{-1}$. The volume was adjusted to 50 µl with $H_2O$, briefly mixed then incubated overnight at room temperature.
200 µl of herring sperm DNA (50 µg ml$^{-1}$) was added to the labelling solution then it was phenol extracted and ethanol precipitated. The DNA was redissolved in 100 µl of T.E and the radioactivity was measured by Cerenkov counting in a scintillation counter. Specific activities exceeding 4 x $10^8$ cpm/µg were usually obtained.

### 2.4.11 b) Oligonucleotide labelling by phosphorylation

The following were mixed together: 100 ng of oligonucleotide, 2.5 µl 10 x kinase buffer, 5 µl (50 µCi) [$\gamma$-$^{32}$P] ATP at 3000Ci mmol$^{-1}$, 1µl (10 units) of T4 polynucleotide kinase, $H_2O$ to 25 µl. The reaction was mixed then incubated at 37$^0$C for 1 hr. After labelling, 200 µl of herring sperm DNA (50 µg ml$^{-1}$) was added and the DNA was then phenol:chloroform extracted and ethanol precipitated twice in the presence of 2 M ammonium acetate. This removed most of the unincorporated label.

### 2.4.12 Southern blotting

DNA fragments from an agarose gel were transferred either to Hybond-N (Amersham) or Genescreen (DuPont) filters, both of which are positively charged and are nylon based. The same blotting apparatus was used for both filters. The original method for this transfer was devised by Southern (1975).
For Southern transfer with Hybond-N the gel, after electrophoresis, was soaked in 500 ml 0.25 M HCl for 2 x 12 mins with gentle shaking. The gel was then briefly rinsed in $dH_2O$ and then soaked in 0.4M NaOH, 0.6 M NaCl for 2 x 12 mins. This solution is also used as the transfer buffer.
For Southern transfer with Genescreen the gel, after electrophoresis, was soaked in Genescreen transfer buffer with gentle shaking for 30 mins.

Blotting apparatus

A glass plate was balanced on the sides of a sandwich box. A length of Whatman 3MM paper (e.g 15 cm x 40 cm) was draped over the glass plate and into the box. The box was then half filled with transfer buffer and the paper was allowed to become wet. On to this the gel was placed and air bubbles which formed were gently removed. A piece of membrane (Hybond or Genescreen), cut to the same size as the gel was soaked in $H_2O$ for at least 30 mins then in transfer buffer for 10 mins. This membrane was then placed on top of the gel and again air bubbles were gently removed. Three pieces of Whatman 3MM paper (cut to size of the gel) were placed on top of the membrane. On to this was then added a thick wad (10 cm) of blotting paper which was again cut to the size of the gel. A glass plate was then placed on top of the paper and then a weight (500 g) was placed on to the glass plate. This was left at room temperature overnight.

Washing and fixing of membranes

After transfer Hybond membranes were washed in 200 ml 2 x SSC for about 5 mins and then allowed to dry at room temperature for at least an hour. The DNA was then fixed to the membrane by wrapping it in Saran wrap and placing it in a Hybaid crosslinker. It was then exposed to 0.4 $J/cm^{-2}$ of U.V light and the filter was removed. Genescreen membranes were simply allowed to dry at room temperature after transfer.

### 2.4.13 Hybridisation and post hybridisation washes

**a) Hybond N**

**i) Using random labelled probes**

Hybond N filters were prehybridised at 65°C for at least 4 hrs, usually overnight, in about 50 ml of prehybridisation solution. Prehybridisatiocn of membranes ocurred in sealed plastic bags.

For hybridisation the membrane was transferred to another bag. The probe and 5 mg of herring sperm DNA were denatured by boiling for 10 mins and added immediately to 10 ml of prehybridisation solution at 65°C. This was then added to the plastic bag containing the membrane and sealed. The bag was incubated overnight at 65°C with shaking.

Washing

After hybridisation the membrane was washed with the following solutions.

1st wash - 2 x SSC, 0.1 % SDS, 30 mins at room temperature
2nd wash - 0.1 x SSC, " " " "
3rd wash - " " " 30 mins at 37°C
4th wash - " " " 30 mins at 65°C
Each membrane was washed in about 200 ml of the wash solution.

**ii) Using oligonucleotides**

Oligonucleotide hybridisations were performed exclusively on Hybond N membranes. The membrane was prehybridised in 30 ml of oligonucleotide prehybridisation solution at 65°C overnight in a shaking water bath.
After the membrane was prehybridised, the probe was added to 10 ml of fresh prehybridisation solution and the membrane sealed into a new bag. Hybridisation took place overnight at 42°C.

Washing

After hybridisation the membrane was washed in the following solutions:

1st wash - 6 x SSC, 0.1 % SDS, 30 mins at 42°C
2nd wash - 2 x SSC, " "
3rd wash - 0.1 x SSC, " "

**2.4.13 b) Genescreen**

Prehybridisation for Genescreen was done for at least 30 mins in about 50 ml of Genescreen prehybridisation solution. This occurred in a shaking water bath at 65°C. For hybridisation the boiled probe was added to 10 ml of prehybridisation solution and added to a sealed plastic bag along with the memebrane. This was incubated overnight at 65°C with shaking.

Washing

After hybridisation the membrane was washed in the following solutions:
1st wash - 2 x SSC, 15 mins at room temperature
2nd wash - 0.1 x SSC, 1 % SDS, 15 mins at 65°C
3rd wash - " " "

### 2.4.14 Colony blotting

This method was based on the procedure used by Grunstein and Hogness (1975). It involved growing bacteria on the surface of the circular membrane (Hybond-N) which was laid onto the agar plates. Master plates were made by pre-wetting a fresh membrane on an agar plate and then placing on a piece of sterile 3MM paper. The membrane with the bacterial colonies was then pressed against the fresh membrane and aligned with needle marks. The membranes were then placed on fresh L agar/ amp plates and grown for at least six hours. Master plates were then wrapped in parafilm and stored at 4°C.

After bacterial growth, the membrane was removed with sterile forceps and placed colony side up onto 3 sheets of Whatman 3MM paper which was soaked with denaturing solution, and left for 7 mins.
The membrane was then placed onto fresh 3MM paper which was soaked with neutralising solution and left for 3 mins. This was then repeated with fresh neutralising solution. The membrane was then washed in 2 x SSC and the bacterial debris gently removed with a gloved hand. Membranes were then allowed to dry completely then the DNA was fixed to the membrane as previously described.
Hybridisation and washing procedures were carried out as for Hybond.

### 2.4.15 Developing autoradiographs

Autoradiographs were developed in a dark room by washing the x-ray film for 1 min in x-ray developer (Agfa-Gevaert) in a tray. The developing reaction was stopped by washing the film for 30 secs in a 3 % glacial acetic acid solution. The film was then fixed in X-ray fixing solution (Agfa-Gevaert) for 2 mins. Finally, the film was washed in cold running water for at least 10 mins and then dried in a drying cabinet.

### 2.4.16 DNA sequencing

The chain termination method of DNA sequencing was used throughout and was based on the original method of Sanger and Coulson (1977).

**a) Double stranded sequencing**

i) Preparation of template

A single fresh colony of bacteria transformed with the plasmid of interest was used to inoculate 3 ml of L broth containing the appropriate antibiotic. This was grown at 37$^{0}$C with vigourous shaking for at least six hrs. The bacteria were pelleted by centrifugation for 1 min in a microfuge at top speed.
Double stranded DNA was prepared as described for the magic mini-prep method. Alternatively DNA from a CsCl preparation was used.

ii) Denaturation of double stranded DNA

3-5 μg of DNA was denatured by adding 0.1 vol of 2 M NaOH, 2 mM EDTA and incubating at 37°C for 30 mins. 3 vols of ice cold ethanol and 0.1 vol of 3 M sodium acetate was added, mixed and stored at -70°C for 15 mins or for 30 mins at -20°C. The DNA was recovered by centrifugation in a microfuge for 10 mins at top speed.

iii) Annealing of primer to template

The denatured DNA was resuspended in 7 μl $H_2O$, 2 μl Sequenase (United States Biochemical Corporation) reaction buffer and 1 μl primer (1 pmol $\mu l^{-1}$). This was mixed then heated to 65°C for 5 mins. The heating block containing the tube was then removed and allowed to cool to below 35°C. The annealed template was then briefly spun and stored on ice.

iv) Labelling reaction

Still on ice, the following components were added to the annealed template-primer; 1 μl 0.1 M DTT, 2 μl 1:15 diluted labelling mix, 0.5 μl [ $\alpha$-$^{35}$S] dATP (10 μCi $\mu l^{-1}$) and 2 μl 1:8 diluted Sequenase. 1 μl of Mn buffer was added when sequence close to the primer was needed. This was mixed and left at room temperature for 3-5 mins.

iv) Termination reaction

The reaction was terminated by transferring 3.5 μl of the labelling reaction into 2.5 μl of each of the ddNTP solutions (ddGTP, ddTTP, ddATP, ddCTP), mixing and incubating at 37°C for 3-5 mins. 4 μl of stop solution, was added to each

tube. Samples were heated at 80°C for 5 mins before being loaded onto the gel. Alternatively, they were stored at - 20°C until needed.

**2.4.16 b) Single stranded DNA sequencing**

i) Preparation of an infected culture

A single fresh colony containing the phagemid pT7T3 18 or 19u, was used to inoculate 3 ml of L broth containing ampicillin. Helper phage M13 K07 was added to a final concentration of 2 x $10^7$ pfu/ml and incubated for 1.5 hrs at 37°C with vigourous shaking. Kanamycin was added to a final concentration of 70 μg $ml^{-1}$ and the incubation was continued for 14 -18 hrs at 37°C, again with shaking.

ii) Preparation of single stranded template

1.5 ml of the infected culture was centrifuged in a microfuge at top speed for 5 mins. 1.3 ml of supernatant was transferred to a fresh tube and to it was added 200 μl of 20 % polyethylene glycol in 2.5 M NaCl. The tube was mixed by inverting several times followed by gentle vortexing. This was left to stand at room temperature for 30 mins. The precipitated bacteriophage were recovered by centrifugation for 5 mins in a microfuge at top speed. All of the supernatant was discarded without disturbing the pellet. 100 μl of T.E was used to resuspend the pellet and then this was phenol / chloroform extracted, then chloroform extracted. After being centrifuged in a microcentrifuge for 2 mins at top speed the DNA, from the aqueous layer, was ethanol precipitated and the pellet dissolved in 50 μl $dH_2O$.

5 μl (~1μg) of this was usually used for a sequencing reaction.

iii) Annealing of template to primer

5 µl of DNA, 2 µl $H_2O$, 2 µl sequenase reaction buffer and 1 µl of primer (1pmol $\mu l^{-1}$) were mixed and annealed in the same fashion as for double stranded DNA. Single stranded DNA was then sequenced as described for double stranded DNA.

**2.4.16 c) Polyacrylamide sequencing gels**

6 % acrylamide solution was made by mixing 5.7 g acrylamide, 0.3 g bis-acrylamide, 48 g urea, 1.1 g Tris, 0.55 g orthoboric acid and 600 µl 0.5M EDTA. This was made up to 100 ml with $dH_2O$ and mixed until it dissolved.
30 ml of this solution was used for a normal size gel (35 cm). 0.5 ml of 10 % ammonium persulphate was added and polymerisation was initiated by the addition of 25 µl TEMED (N,N,N',N' - Tetramethylethylene - diamine). For the large gels (65 cm) the volumes described above were doubled.
The solution was poured between the glass plates and allowed to set for at least 1hr. These gels were usually pre-warmed for at least 30 min at 1800 volts before the addition of the sample. 1 x TBE was used as the buffer and samples were usually run at 1500 volts for the appropriate length of time.

After electrophoresis the gel was soaked in 10 % glacial acetic acid for 30 mins to remove the urea then transferred to a sheet of Whatman 3MM paper. The gel was then dried for 30 mins at 80°C under vacuum.

## 2.5 Growth of mammalian tissue culture cells

Neuro2a cells (murine neuroblastoma)

These cells (from ICN-FLOW) were grown in Dulbecco's modified Eagle medium (DMEM) which was supplemented with 10 % foetal calf serum. Penicillin and streptomycin (100 $\mu g\ ml^{-1}$) were also added to this medium as a precautionary factor against bacterial infection. Cells were incubated at 37°C in an atmosphere of 6 % $CO_2$ ,in air, until 80 % confluency was reached. The cells were then removed from the plate by scraping, either with a cell scraper or a cell lifter, and split 1:8. This essentially involves a 1:8 dilution of cells with fresh media. Freshly split cells were then returned to the incubator. For transfections, the cells were plated out in 60 $mm^2$ dishes (6 ml) at 1-2 x $10^5$ cells $ml^{-1}$ as counted in a haemocytometer. Transfections were performed the next day.

## 2.6 Storing of tissue culture cells

When cells were approaching confluency they were removed from the plates as described above. The cells were recovered by centrifugation for 3 mins at 1000 x g 4°C. The cell pellet was resuspended in 10 ml of freezing mix per 75 $cm^2$ flask and centrifuged again. The supernatant was decanted and replaced with 5 ml of fresh freezing mix. The cells were resuspended and aliquoted out into 0.5 ml amounts in Nunc screw top freezing vials. These were stored at -70°C overnight before being transferred into liquid nitrogen.

For thawing, vials were removed from the liquid nitrogen and quickly placed into a 37°C water bath. Once thawed, the cells the cells were decanted into 5 ml of fresh media at 37°C and collected by centrifugation for 10 mins at 1000 x g. The cells were resuspended in 10 ml of fresh media and plated out into a 25 $cm^2$ flask.

## Chapter 3 Construction and screening of a cosmid genomic library

### 3.1 Introduction

Analyses of PrP genes from Syrian hamster (Basler *et al.*, 1986), human (Puckett *et al,* 1991) and mouse (Westaway *et al.*, 1987) have shown that they all have a similar gene structure. A single exon which contains the entire open reading frame is separated from 1 or 2 upstream exon (s) by intron (s) of at least 10 kb in length. Depending on the species there are 1 or 2 non-coding upstream exons. Mouse and sheep have 2 upstream exons (Westaway *et al.*, 1991, 1994 b), whereas hamster and human have only been shown to have 1 (Basler *et al.*, 1986; Puckett *et al.*, 1991). The biological functions of these upstream exons are not known but they may play a regulatory role.

Work on other genes has shown that regulatory elements can be found throughout the gene: both in 5' and 3' flanking regions and also in introns. 5' and 3' flanking regions can be defined as non-coding genomic sequences which are found 5' to the first exon and 3' to the last exon, respectively. There are many examples of regulatory elements in 5' flanking regions of genes. Transcriptional promoters are usually found in the 5' flanking regions of genes. Izumi *et al.*, (1992) have analysed the promoter region of the Alzheimer's disease amyloid precursor protein gene. They found that both positive and negative regulatory elements are present in the promoter region and reside within several base pairs of each other. Others have shown that tissue specific expression can be due to elements in the 5' flanking region. For example, in the human neurofilament light chain promoter a 210 bp conserved region has been shown to be required for neuronal specific expression (Yazdanbakhsh *et al.*, 1993). Work on the human α-globin has demonstrated that sequences 30-50 kb upstream of the first exon contain critical regulatory elements (Higgs *et al.*, 1990). In addition to 5' flanking regulatory elements there are also some examples of regulatory elements in introns e.g in the first intron of the α-I(1) collagen gene (Bornstein and McKay,

EDINBURGH UNIV. LIBRARY

1988) and in 3' regions e.g silencer and enhancer elements of the duck α-globin gene ( Targa *et al.*, 1993).

The regions of the murine PrP gene which may be important for the regulation of gene expression are not known. To address this, one must obtain a genomic clone of the PrP gene. Such a clone would preferably contain 5' flanking sequences, the non-coding exons, and the introns of the PrP gene.

There are a variety of vectors which may be used in the construction of genomic libraries: plasmids, lambda, P1, yeast artificial chromosome (YACs) and cosmids. Each of these vectors can accommodate different insert sizes. YACs have the capacity for propagating the largest inserts; ~ 400 kb (Burke *et al.*, 1984) whereas the maximal insert size for most plasmids is approximately 15 kb. The maximal insert sizes for λ and cosmid vectors are approximately 23 and 45 kb respectively. The objective of this project was to obtain a single clone of the entire mouse PrP gene, together with its flanking sequences. Obtaining a single clone, as opposed to several overlapping clones, would facilitate subsequent experiments such as mapping and characterising the promoter region as well as in the generation of mice carrying functional PrP transgenes and deletions of such transgenes. Such transgenic mice could be used to study developmental and tissue specific expression of the PrP gene. As mentioned previously, such regulatory regions may lie many kilobases away and therefore it is crucial to obtain a single clone which contains maximal 5' and 3' flanking sequences.

Westaway *et al.*, (1987) had previously shown, by nucleotide sequencing and Southern blotting, that the exon containing the open reading frame in the mouse PrP gene was approximately 2 kb and that this exon was separated from an upstream exon by at least 11.5 kb of intron sequence. At the time this study was initiated only one untranslated 5' exon had been identified (Westaway *et al.*, 1987).

To clone PrP by cosmid cloning therefore presented the best approach as one had a better chance of obtaining the full length gene in one clone. The cosmid vector used was pWE 15 (Wahl *et al.*, 1987) and like other cosmid vectors it can accept and propagate fragments in the size range of 35-45 kb. Cosmids are essentially hybrids between plasmid and lambda vectors. They are designed around the fact that the enzymes which package viral DNA into the phage head, during normal viral replication, recognise specialised sites called *cos* sites. These *cos* sites are separated from each other by 37-52 kb of viral DNA. Cosmids utilise these *cos* sites so that 35-45 kb of insert DNA can be ligated to the vector and efficiently packaged *in vitro* into phage heads. These particles can now be used to infect E.coli but they do not produce plaques as the normal virus would, this is because most of the viral DNA has been replaced with insert DNA. Instead, antibiotic resistant bacterial colonies are produced. These can then be screened by hybridisation to determine the clone of interest (see fig 5).

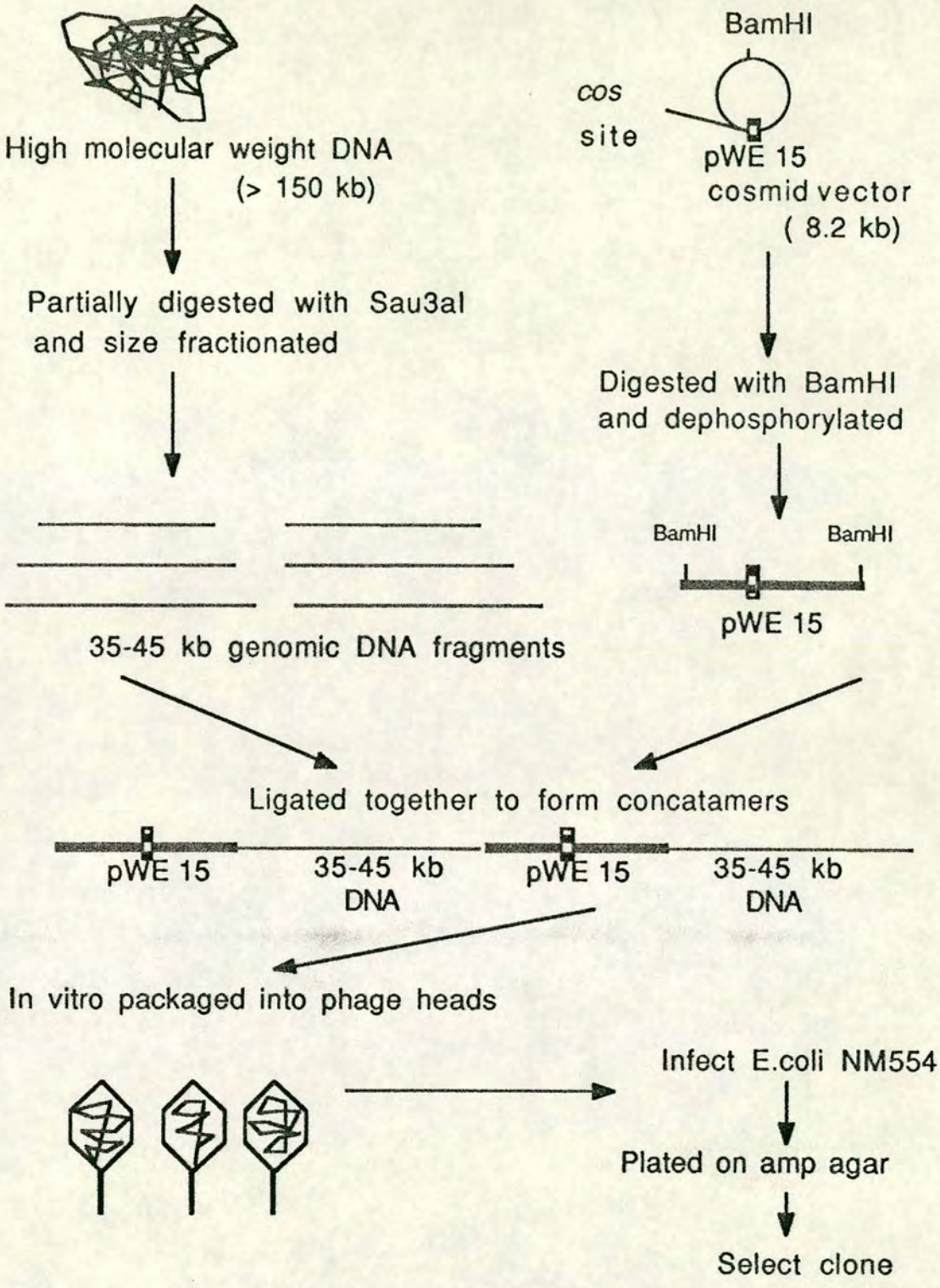



**Figure 5.**

Schematic representation of the steps involved in the construction and screening of a cosmid genomic library. High molecular weight mouse DNA (129/Ola) was isolated and partially digested with Sau3AI. The DNA was size fractionated on a sucrose gradient to give DNA in the 35-45 kb range. The size selected DNA was then ligated to the dephosphorylated pWE 15 cosmid vector to produce concatamers. These concatamers were then packaged into phage heads by *in vitro* packaging. The recombinant phage particles then infected E.coli and ampicillin resistant colonies were produced. This resulted in the production of a genomic library which was then screened for a PrP cosmid clone.

### 3.2 Preparation of vector DNA

20 μg of pWE15 cosmid vector (Stratagene) was digested to completion with BamHI. The DNA was then dephosphorylated to prevent self ligation. This was followed by a phenol:chloroform extraction and an ethanol precipitation. The DNA was redissolved in 20 μl T.E (pH 8).

### 3.3 Isolation of high molecular weight DNA

A 162 $mm^2$ flask containing murine embryonic stem cells (ES), grown to confluency, was used as the source of DNA. These cells were derived from a 129/O1a mouse. PrP null mice had previously been generated with these cells by homologous recombination (Manson *et al.*, 1994) and therefore it seemed prudent to generate future transgenic mice from the same genetic background. 129/O1a mice are Sinc s7 homozygous. Sinc is the major gene which controls the incubation period of scrapie in mice (see section 1.4.1) and mice which are s7 homozygotes generally have shorter incubation periods, with most scrapie strains, than their p7 homozygous counterparts (Bruce *et al.*, 1991). This mouse genotype was therefore also chosen as it is the genotype in which disease generally manifests itself more rapidly.
To the flask was added 10 ml of a solution containing TEN 9, proteinase K (500 μg $ml^{-1}$) and 1 % SDS. This was left shaking overnight at 37°C. 10 ml of phenol was then added and left gently shaking at room temperature for 1 hr. This solution was then decanted into a 50 ml centrifuge tube and spun for 10 mins at 1000 x g. The aqueous phase was then phenol extracted and then chloroform extracted. The aqueous phase was then dialysed against 4 L of T.E (pH 8) overnight at 4°C. The DNA concentration was determined and an aliquot was examined by 0.3 % agarose gel electrophoresis (on top of 1 % support gel). By comparison to Gibco BRL high molecular weight markers, the purified DNA was estimated to be in the 150-200 kb range.

### 3.4 Partial digestion of high molecular weight DNA using Sau3AI

The high molecular weight DNA was handled with cut off pipette tips (all sizes) to prevent shearing. Conditions were established which allowed partial digestion of the high molecular weight DNA with Sau3AI to generate fragments of DNA in the appropriate size range for insertion into the cosmid vector (35-45 kb). BamHI is the unique cloning site of the pWE 15 cosmid vector. However, Sau3AI was chosen to digest the genomic DNA as it digests DNA more frequently than BamHI. This is because it has a 4 bp recognition site. It also generates compatible ends with BamHI digested DNA. The following conditions resulted in digested DNA in the correct size range. 4 µg of DNA was gently mixed with 10 µl the appropriate restriction enzyme buffer and 86 µl $H_2O$. This was allowed to stand at $4^0$C for several hours to ensure that a homogenous DNA solution was formed before the restriction enzyme was added. 1 µl of Sau3AI (0.05 units $\mu l^{-1}$) was added to the pre-warmed DNA solution and incubated at $37^0$C for 30 mins. The restriction enzyme was inactivated by heating the mixture to $75^0$C for 15 mins. This reaction was scaled up to give 260 µg of digested DNA in total. The samples were pooled, phenol:chloroform extracted and ethanol precipitated. The DNA was resuspended in 200 µl T.E pH 8. 10 µl of this was then run on a 0.3 % agarose gel to ensure DNA was of the correct size.

### 3.5 Size fractionation of DNA using sucrose/salt gradient centrifugation

#### a) Preparation of gradient

A 10-40 % continuous gradient was prepared in a SW-28.1 centrifuge tube using the following solutions:

40 % sucrose, 1 x T.E, 1M NaCl

30 % sucrose, 1 x T.E, 1M NaCl

20 % sucrose, 1 x T.E, 1M NaCl

10 % sucrose, 1 x T.E, 1M NaCl

2.5 ml of the 40 % solution was added to the tube. On top of this 2.5 ml of the 30 % solution was gently added. The same was done with both the 20 % and the 10 % solutions. The gradient was then prepared by centrifuging at 90 000 x g for 3 hrs at $12^0$C .

### 3.5 b) Loading DNA sample and centrifugation

Approximately 260 µg of partially digested DNA in a 200 µl volume of T.E was gently layered onto the gradient. The sample was then centrifuged at 90 000 x g for 18 hrs at $12^0$C.

### c) Collection and electrophoresis of fractions

500 µl fractions were collected from the bottom of the tube after centrifugation. 35 µl of each fraction was run on a 0.3 % agarose gel overnight at 15 V. High molecular weight markers were used for size comparison. Samples 4, 5, 6, 7 ,8, and 9 were pooled as they contained the DNA in the correct size range. The sucrose concentration was then reduced to below 10 % by diluting the pooled fractions five fold in $H_2O$. The DNA solution was then ethanol precipitated and then DNA recovered by centrifugation at 8000 x g for 30 mins at 4°C. The DNA was resuspended in 100 µl T.E (pH 8).

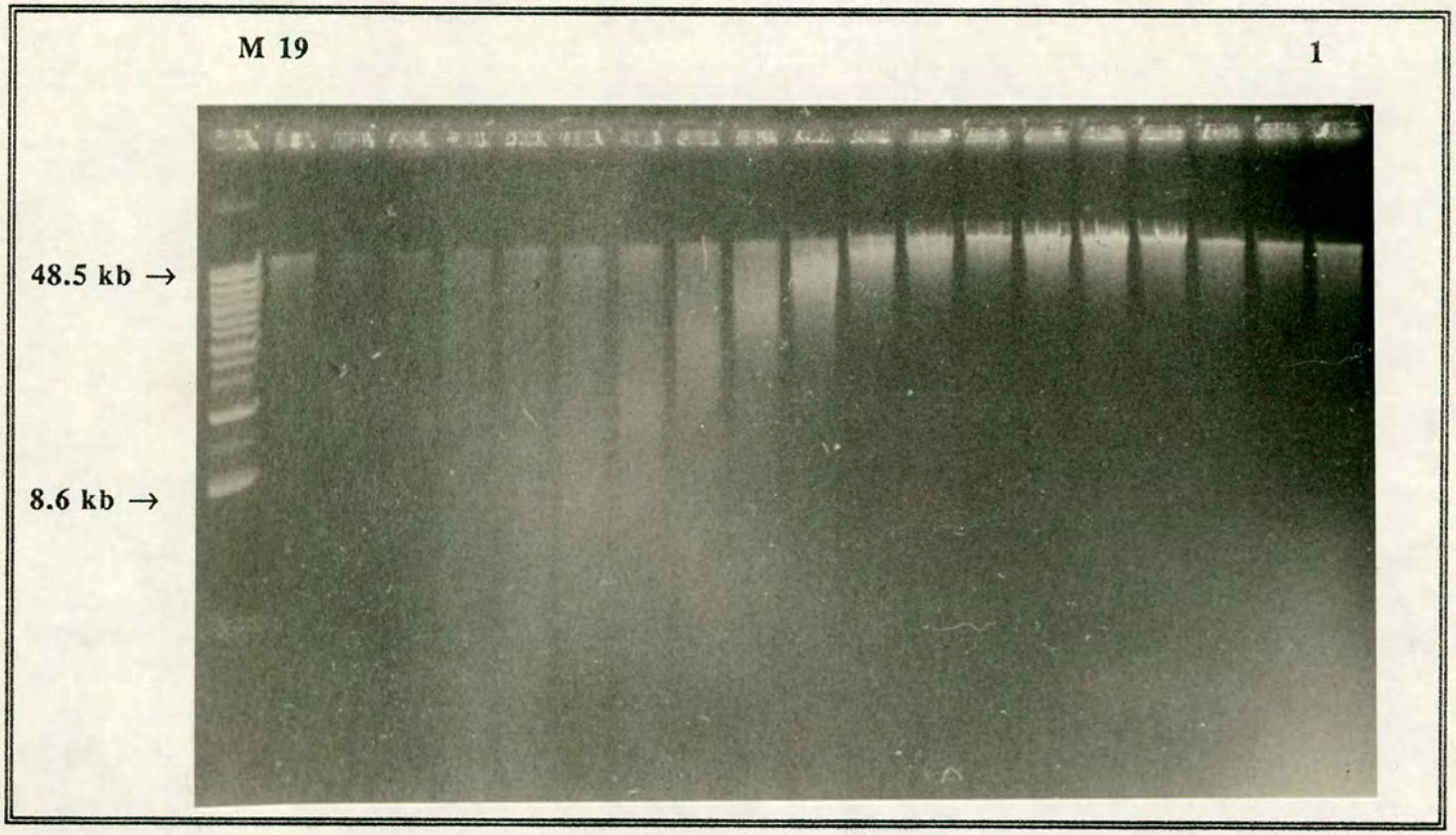



**Figure 6.**

Size fractionation of Sau3AI partially digested mouse genomic DNA.

After optimal conditions were established for DNA digestion (see 3.4), the digested DNA was size fractionated to remove smaller species of DNA which would otherwise result in inefficient *in vitro* packaging. Fraction 1 represents the fraction at the bottom of the tube wereas fraction 19 was obtained from the top of the tube. 35 µl (out of a total of 500 µl) of each fraction was run on a 0.3 % agarose gel overnight at 15 V . Fractions are numbered 1-19 and high molecular weight markers (Gibco BRL) are designated as M. The markers range from 48.5 kb to 8.2 kb (only some are arrowed). Samples 4, 5, 6, 7, 8 and 9 represented DNA of approximately 45 kb and were chosen for the ligation stage.

### 3.6 Ligation and packaging of DNA

#### a) Ligation of vector to size selected DNA

DNA was ligated the addition of approximately 1.5 µg size fractionated DNA, 3 µg of dephosphorylated BamHI digested pWE 15, 4 µl 5 x ligation buffer, 1 µl T4 DNA ligase, made up to 20 µl with $H_2O$. This was incubated overnight at 16°C.

#### b) *in vitro* packaging of DNA

Gigapack II gold packaging extracts (Stratagene) were used.
A freeze/thaw extract and a sonic extract were removed from the freezer and the freeze/thaw mix was quickly thawed. Four microlitres of the ligated DNA was then mixed in and placed on ice. Fifteen microlitres of the sonic extract was added to the freeze/thaw extract containing the DNA and mixed with a pipette tip. The tube was then briefly spun to ensure that the contents were at the bottom of the tube. Incubation for 2 hrs at room temperature then took place. After this 500 µl of the phage dilution/SM buffer was added. Twenty microlitres of chloroform was added and the solution was gently mixed. Debris was removed after a brief spin and the supernatant was stored at 4°C.

**c) Determining the titre of the cosmid library**

A single fresh colony of E.coli NM554, grown at 37°C overnight on L agar was used to inoculate 25 ml of L broth containing 0.2 % maltose. This was incubated at 37°C overnight with shaking. The bacteria were pelleted by centrifugation at 1000 x g for 10 mins. The pellet was resuspended in 12.5 ml of 10 mM $MgSO_4$. More solution was gradually added until the $OD_{600nm}$ was adjusted to approximately 0.5. The packaged DNA was diluted 1:10 and 1:50 in SM buffer. Twenty five microlitres of each dilution was mixed with 25 µl of the diluted E.coli NM554 and incubated at room temperature for 30 mins. 200 µl of L broth was then added and incubated with intermittent shaking at 37°C for 1 hr. The bacteria were pelleted by centrifugation in a microfuge at 11600 x g for 30 secs and resuspended in 50 µl of fresh L broth. Each dilution was spread out onto L agar plates containing ampicillin (50 µg $ml^{-1}$ ) and incubated overnight at 37°C.

The colonies were then counted and the titer calculated to be 2.2 x $10^5$ transformants $\mu g^{-1}$ of size fractionated DNA. The titre is between the expected values of 5 x $10^4$ and 5 x $10^5$ transformants $\mu g^{-1}$ of size fractionated DNA (Sambrook *et al.*, 1989). There was approximately 1.5 µg of size fractionated DNA in the ligation reaction. This meant that on packaging and plating of the entire ligation mix there should be approximately 3.3 x $10^5$ colonies. After calculating the titre the rest of the ligation was packaged in a similar manner (4 µl per packaging mix).

The number of colonies which one must screen in order to have a good chance of obtaining a specific clone is dependent on the average insert size and the total size of the genome from which the library is made. If one assumes that the average insert size was 40 kb, that the size of the mouse genome is 3 x $10^6$ kb and that the number of colonies screened was 3.3 x $10^5$ then according to the formula of Clarke and Carbon (1978) there was a > 95 % probability of obtaining a positive clone.

### 3.7 Plating, replica plating and storing of the cosmid library

The packaged DNA (at ~ 30000 colonies equivalent) was mixed with an equal volume of E.coli NM554 at $OD_{600nm}$ = 0.5. This was incubated at 37°C for 30 mins. Four volumes of L broth were added to each of the tubes and incubated for 1 hr at 37°C with intermittent shaking. The culture was centrifuged at 1000 x g for 10 mins and then resuspended in 1.5 ml L broth. The culture was then spread onto 132 $mm^2$ circular Hybond-N filters which were placed on top of 150 $mm^2$ L agar/amp plates. Ten plates were produced in total and were incubated overnight at 37°C. According to the cosmid titre there should have been at least 300,000 colonies produced.

A replica of the library was made by pre-wetting a fresh Hybond N membrane on a L agar/ampicillin plate then placing on a piece of sterile 3MM paper. The membrane with the bacterial colonies growing on it's suface was then pressed on to the fresh membrane. The filters were then aligned with three needle marks, replaced onto the L agar/amp plates and re-incubated at 37°C for at least 4 hrs. This was done for all the filters of the library. The master filters were then stored at -70°C by placing on L agar/amp plates containing 25 % glycerol and wrapping in parafilm.

### 3.8 Screening of the cosmid library

Filters were prepared for prehybridisation as described in section 2.4.14. Prehybridisation and hybridisation also took place using the conditions described in section 2.4.13.

The filters were screened for a PrP positive cosmid clone by using a 3.6 kb XbaI fragment which covered the entire length of exon 3. This probe was called probe X and was derived from pUC N 18.

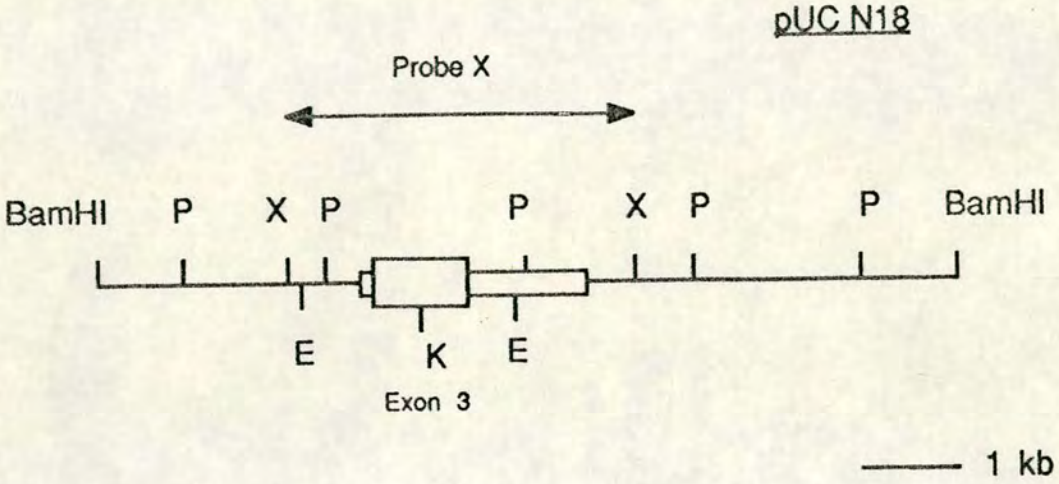



**Figure 7.**

pUC N18 is a 8.8 kb genomic clone of a NZW mouse PrP gene (provided by D. Westaway). This clone encompasses exon 3 and surrounding region. The entire open reading frame is indicated by a large open box. The narrower open box designates the untranslated regions of exon 3. The single straight line at the 5' end of the insert represents part of the large intron whereas the same line at the 3' end represents 3' flanking sequences. Restriction sites: P, PstI; E, EcoRI; K, KpnI and X, XbaI.

The 3.6 kb XbaI fragment was radioactively labelled by random primer extension to approximately 1 x $10^8$ cpm $\mu g^{-1}$ as described in section 2.4.11 a. Prehybridisation and hybridisation of the replica membrane were carried out overnight at 65°C in 10 ml of prehybridisation solution. Membranes were exposed to X-OMAT X-ray film overnight at -70°C.

### 3.9 Results and discussion

Several hybridisation signals were detected on the autoradiograph. The area on the master plate from which the hybridisation signal was emanating was determined by aligning it to the hybridised filter. Since there was a high density of colonies on the master plate it was necessary to scrape off a region of about 5 mm in diameter surrounding the hybridisation signal using a bacteriological loop. This region was estimated to contain about 20 - 40 colonies This was inoculated into 100 µl of L-broth containing ampicillin.

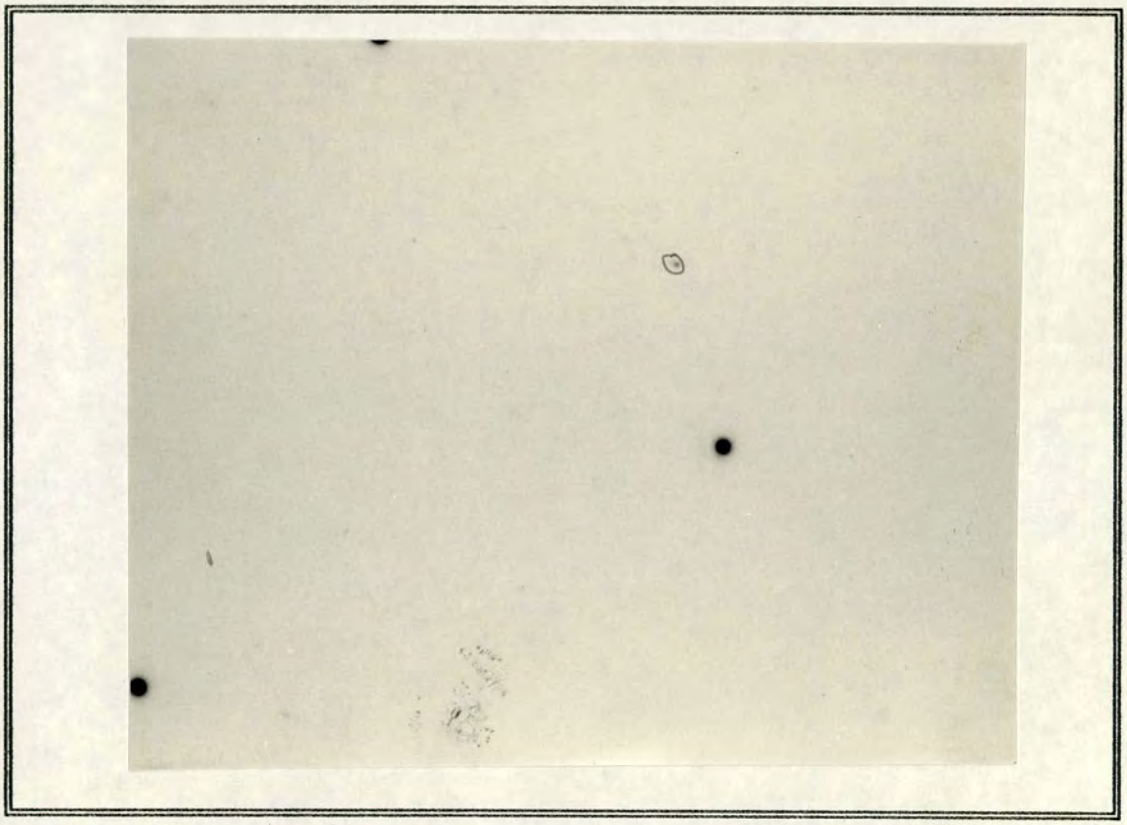


**Figure 8.**

Autoradiograph of primary screen of cosmid library.

Probe X (fig 7) was used to probe the library. After hybridisation the library was autoradiographed at - 70°C overnight. The encircled hybridisation signal is a putative positive, whereas the stronger hybridisation signals surrounding it are alignments made with radioactive ink. This filter represents 1/10th of the total library.

The bacteria were diluted 1:1000 in L broth and 100 µl of this was replated onto a Hybond N filter which was laid on top of a L agar plate containing ampicillin. This was incubated overnight at 37°C. The hybridisation procedure was repeated and the filter autoradiographed.

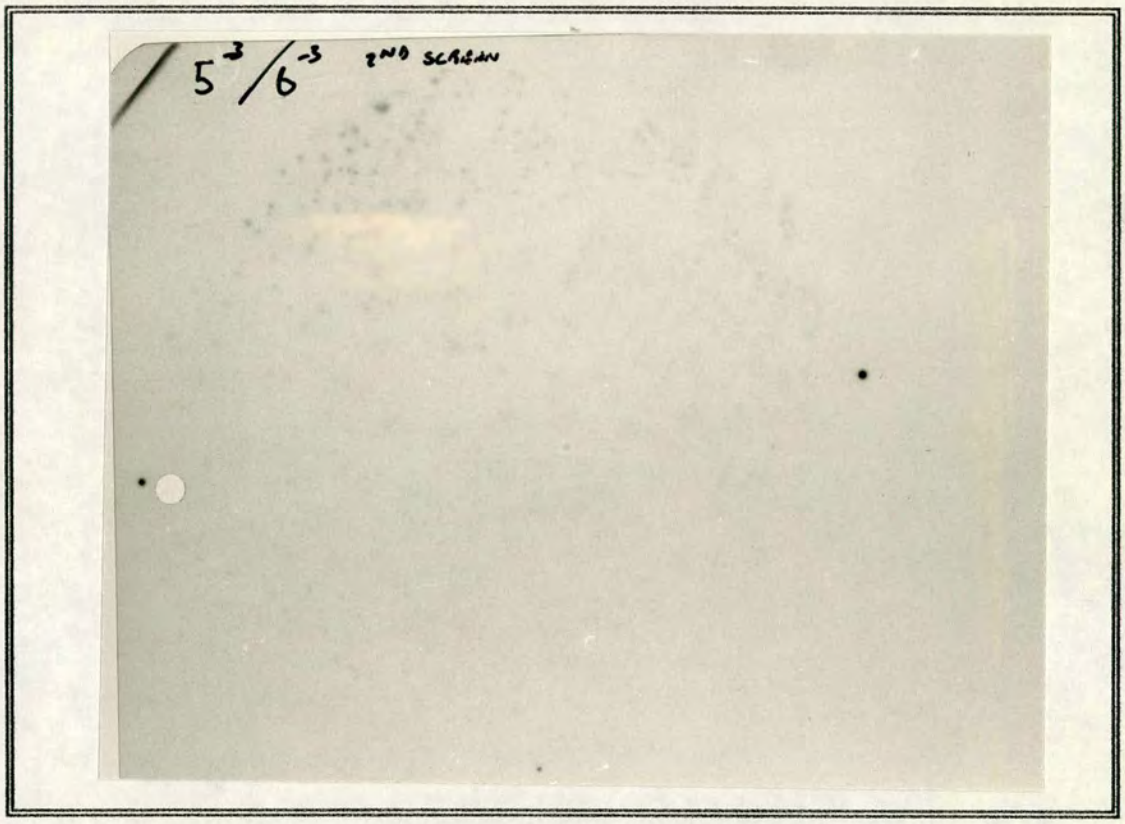



**Figure 9.**

Autoradiograph of secondary screening of putative positive clones.

Again probe X was used in the screening process, and after hybridisation the membranes were autoradiographed overnight at -70°C. Many more hybridisation signals were evident at this stage and are evident at the top left of the photograph. Some of these were picked for another stage of screening. Alignment marks, made with radioactive ink, are seen at the extreme left and right of the figure.

Individual colonies which represented the strongest hybridisation signals were picked. Using a toothpick these were streaked onto a 82 mm gridded Hybond N filter and placed for overnight incubation at 37°C on an L agar plate containing ampicillin. The hybridisation procedure was repeated again.


**Figure 10.**

Tertiary screening of the cosmid library.

This was the last in the screening stages. Colonies from the secondary screening stage which gave the strongest hybridisation signals were plated onto a gridded Hybond N membrane and again prepared for colony hybridisation. Probe X was used and after hybridisation the membrane was autoradiographed overnight at - 70°C. Colonies are numbered from left to right. 50 colonies were screened and 11 of these gave strong hybridisation signals. These were numbers 1, 4, 5, 7, 11, 18, 23, 24, 26, 27, and 30.

The colonies which gave the strongest hybridisation signals were grown up in liquid culture for small scale plasmid preparation. These clones were then subjected to restriction enzyme digestion. Out of the 11 clones selected 7 gave the same restriction pattern when digested with BamHI. 3 clones gave separate and different restriction patterns. No DNA could be isolated from one colony. This gel was blotted onto Hybond-N and again probed with probe X (see fig. 7).

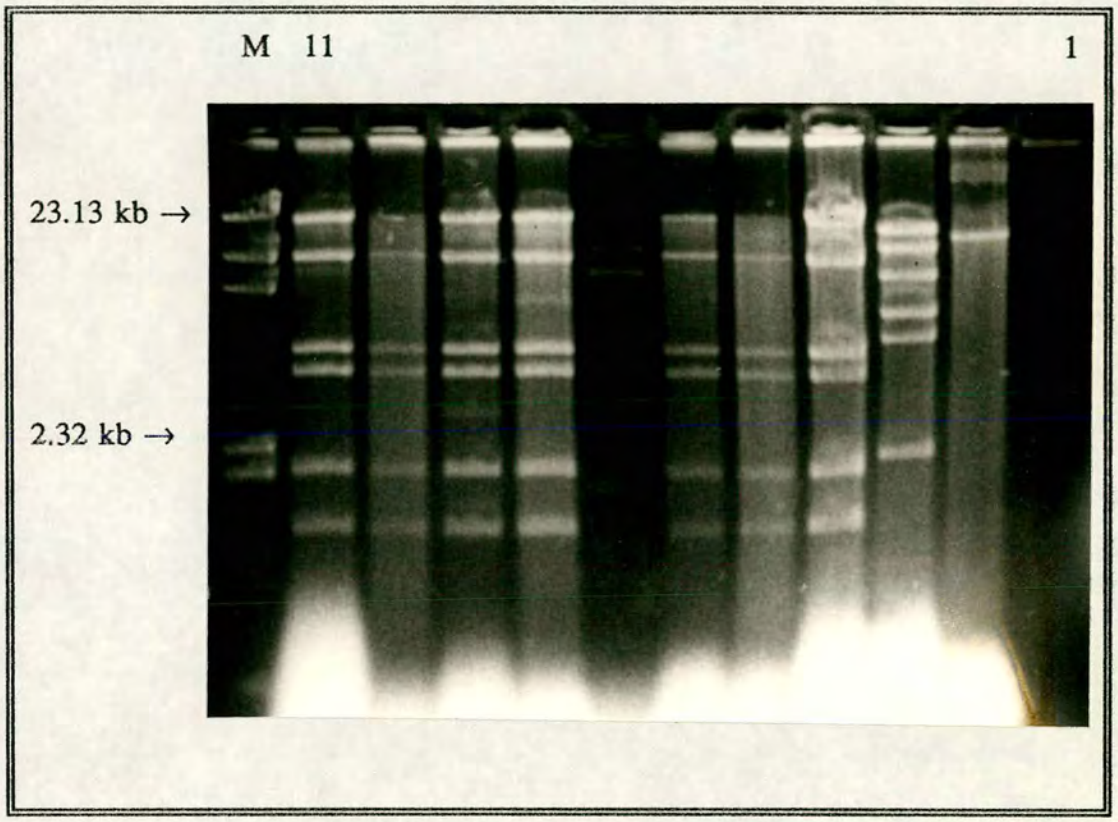



**Figure 11 a.**

a) Gel electrophoresis after DNA from putative positives was restriction digested with BamHI. 0.8 % agarose gel electrophoresis analysis of the 11 BamHI digested putative positive clones. The gel was run at 50 V for 4 hrs and then stained with ethidium bromide. M designates markers and these are HindIII digested Lambda DNA. The sizes of these markers range from 23.13 kb to 0.125 kb (only some are arrowed). Seven of these clones appear to be identical and the other 3 are different. Lane one failed to contain DNA.

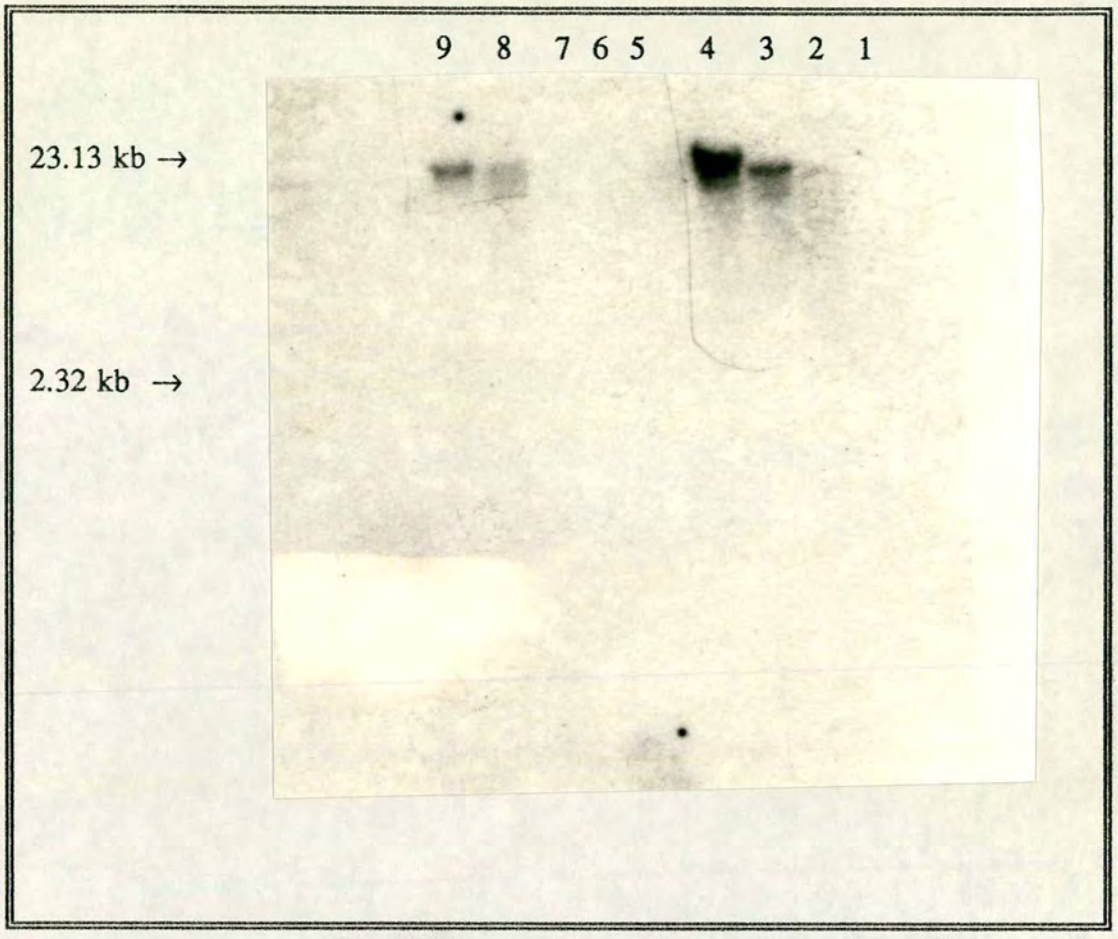



**Figure 11 b.**

b) Southern analysis of the above gel (fig 11 a) when hybridised with probe X.
The membrane was hybridised at 65°C and then washed for 30 mins in i) 2 x SSC, 0.1 % SDS at R/T. ii) 0.1 x SSC, 0.1 % SDS at R/T. iii) 0.1 x SSC, 0.1 % SDS at 37°C. iv) 0.1 x SSC, 0.1 % SDS at 65°C. The filter was then exposed at -70°C for 2 days. Five hybridisation signals were detected. These are seen in lanes 2, 3, 4, 8 and 9. The bands in lanes 4, 8 and 9 are the same size (~ 22 kb). The band in lanes 2 and 3 were estimated to be approximately 18 and 20 kb respectively.

Three of the hybridisation signals came from clones with the same restriction digest pattern (lanes 4, 8 and 9). With these clones the probe hybridised to the same (~22 kb) band. Why the other clones with the same restriction patterns did not result in any hybridisation signal is not clear. Two other clones also gave hybridisation signals (lanes 2 and 3) from smaller sized BamHI fragments of approximately 18 and 20 kb respectively. That three separate clones (lanes 2, 3 and 7 + 8) with distinct restriction patterns all produced hybridisation signals, and that these were all different sizes, suggested that a PrP clone had not been isolated.

Furthermore, none of the hybridised fragments were 8.8 kb in length. Westaway *et al.*, (1987) had previously shown that the exon containing the open reading frame was contained within a 8.8 kb BamHI restriction fragment. It is noteworthy that in these experiments Westaway and his colleagues (1987) used the NZW and IL/N mouse strains. In the experiments described above, the cosmid library was derived from a 129/O1a strain. Others have shown that the 129/O1a mouse strain also has an 8.8 kb BamHI fragment (J. Manson, pers comm) which contains the PrP open reading frame. This allows one, therefore, to discount the possibility that a PrP specific 8.8 kb BamHI fragment was not detected during the hybridisation stages due to differences between mouse strains.

This was confirmed and the specificity of probe X for this 8.8 kb BamHI PrP fragment was checked by genomic Southern hybridisation. This revealed that the 3.6 kb XbaI probe hybridised only to an 8.8 kb restriction fragment when digested with BamHI. This suggested that while the 3.6 kb XbaI probe is specific for PrP, there was still an anomaly regarding cross hybridisation to different clones.

Another replica of the cosmid library was made (see 3.7) and this was screened again. The 3.6 kb XbaI restriction fragment, from exon 3 of the mouse PrP gene, was freshly isolated from pUC N 18 and again used to probe the library (see figure 7), in the same fashion as previously described. This second screening of the cosmid library again resulted in putative positive hybridisation signals. After more extensive screenings (secondary, tertiary and Southern blots of small scale plasmid preparations) these, however were not found to be clones of PrP. While the reasons for failing to obtain a PrP cosmid clone were being investigated (see below) an attempt at constructing a second cosmid library was undertaken. This attempt, however, failed due to persistent problems in obtaining high molecular weight DNA (> 150 kb) from the embryonic stem cells (see 3.3). From 0.3 % agarose gel electrophoresis, the isolated genomic DNA was approximately between 50 - 100 kb and was therefore too small.

Subsequent Southern blot analysis using the digested pWE 15 vector (containing no insert) and the same probe X proved that the probe was cross reacting with the cosmid vector (see fig 12).

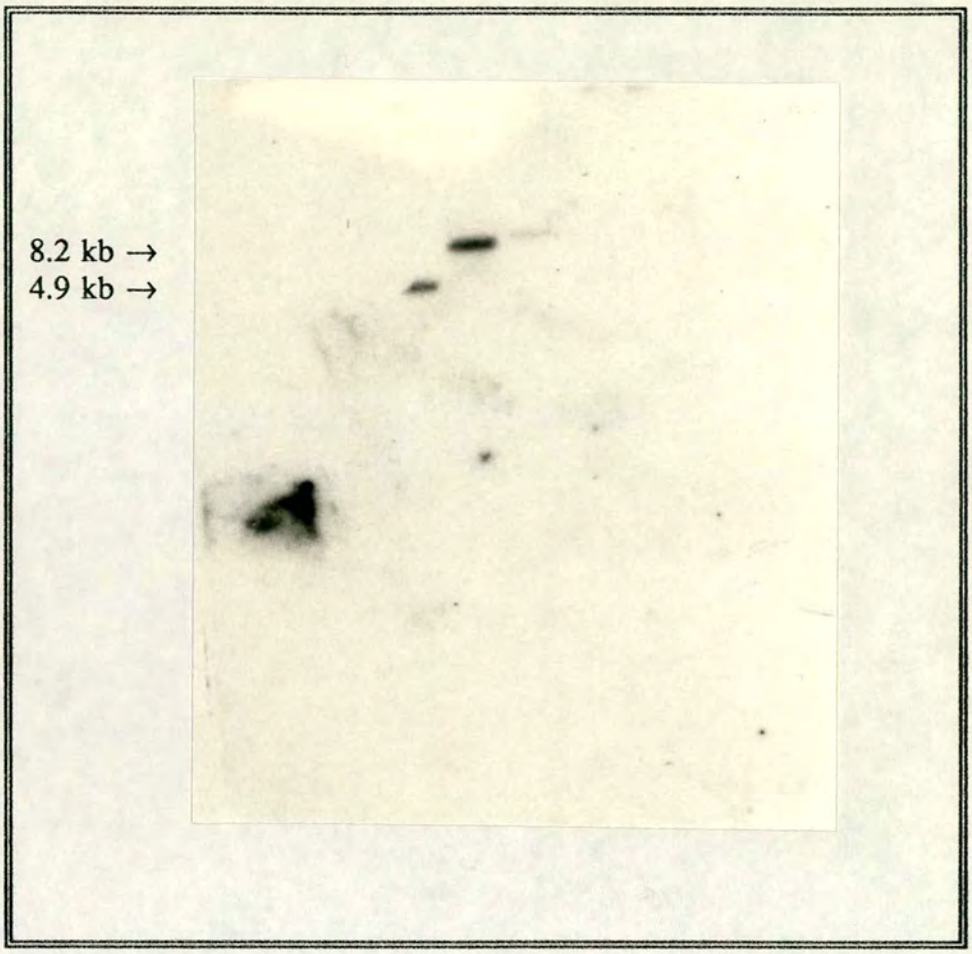



**Figure 12.**

Cross hybridisation of the 3.6 kb XbaI PrP probe to pWE 15 cosmid vector.

pWE 15 was digested with HindIII and BamHI and blotted onto Hybond-N. This blot was then probed with probe X at 65°C and washed in i) 2 x SSC, 0.1 % SDS, at R/T ii) 0.1 x SSC, 0.1 % SDS, at R/T iii) 0.1 x SSC, 0.1 % SDS, at 37°C iv) 0.1x SSC, 0.1 % SDS, at 65°C. All washes were for 30 mins. HindIII digestion produced two bands; 3.3 kb and 4.9 kb, while BamHI digestion produced a 8.2 kb linear. The PrP 3.6 kb XbaI probe hybridised to the 8.2 kb pWE 15 linear band and to the 4.9 kb pWE 15 HindIII band, while not hybridising to the 3.3 kb band, after a 3 day exposure at -70°C.

It therefore seems likely that the 3.6 kb XbaI fragment was not thoroughly isolated from the pUC 18 vector (2.9 kb) during low melting point gel electrophoresis and that some of the vector was subsequently labelled during the random priming step. pWE 15 and pUC 18 share extensive stretches of sequence e.g the ampicillin resistance gene and therefore this may have formed the basis of the cross hybridisation.

It can be concluded therefore that cross hybridisation between the 3.6 kb XbaI probe and pWE 15 vector presented a problem during the screening stage. This could have been avoided if perhaps a shorter probe was used e.g 900 bp KpnI/EcoRI fragment (see fig. 7) as this would have allowed better separation from the vector. Alternatively, the 3.6 kb XbaI fragment could have been cloned into M13 thus preventing any cross hybridisation problems.

That the probe hybridises to vector sequences (see fig. 12) implies that all recombinants in the library should produce a hybridisation signal as all recombinants contain the vector. This did not happen and the reasons why this is so are not clear. However, in a study in 1987, Gibson *et al.*, showed that the insert itself can cause modulation of vector gene expression if it is transcriptionally active. In such circumstances the copy number of the cosmid can be affected, leading to under representation of a particular cosmid clones.

With hindsight it would have been worthwhile to include a positive control during the screening stage. This would have indicated if the probe preparation and the hybridisation conditions were working well. Such a control could have been a probe which would have hybridised to another gene e.g β-globin. In addition, a negative control should also have been incorporated into the experiment. This could have been an E.coli gene such as ompC.

Westaway *et al.*, (1987) have also experienced difficulties when screening both hamster and human genomic libraries for PrP clones. This occurred even when a genome equivalent of recombinants was screened. However, the biochemical nature of this problem was not identified. In the same study, construction and screening of mouse genomic libraries was undertaken.

Successful screening of these libraries was found to be highly dependent on the vector / host system used.

Other workers in our lab have also attempted to use the pWE 15 / NM554 system to clone PrP genes from other species e.g sheep. These attempts have also met with failure in that although clones hybridised, they were subsequently found to have undergone a recombination event. Such an event, therefore, cannot be excluded from the experiments described above and indeed might explain why the expected 8.8 kb BamHI fragment was not found.

# Chapter 4 Isolation, restriction mapping and sequence analysis of the 5' region of the PrP gene.

## 4.1 Introduction

A cosmid clone (designated cos 6.I/Ln J-4) of the murine PrP gene (provided by D. Westaway; see Westaway *et al.*, 1991) was obtained from an I/Ln mouse genomic library. I/Ln mice are *Sinc* p7 homozygous. The scrapie incubation periods of such mice are generally longer than *Sinc* s7 homozygotes when injected with the most commonly used strains of scrapie (Bruce *et al.*, 1991). This means that the s7 homozygous genotype is more commonly associated with disease and as such it would be the preferred genotype in which to study PrP gene regulation.

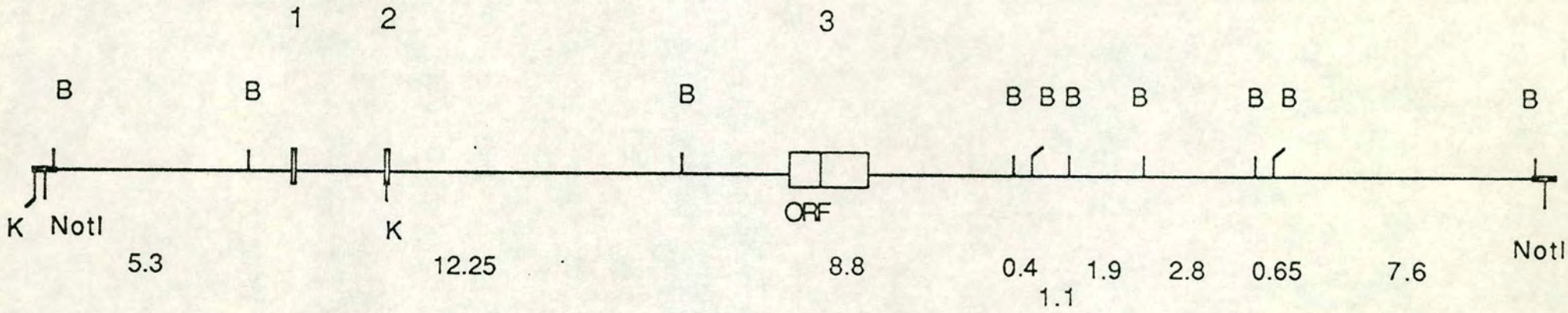
At the beginning of this project there was very little information regarding the structure of the PrP gene (see figure 1, chapter 1). What information there was came from two papers by Basler *et al.*, (1986) and Westaway *et* al., (1987) which described the hamster and mouse PrP genes, respectively. In the 1987 paper it was shown that there was at least one 5' untranslated exon in the mouse PrP gene. This exon was estimated to be greater than 11.5 kb upstream from the open reading frame exon. The existence of introns and measurement of their approximate size was based on several lines of evidence. Southern blots, using oligonucleotides and restriction fragments from various regions of PrP cDNA clones were used to probes genomic blots and blots of genomic clones. The size of the hybridising fragments was then measured. In addition S1 nuclease mapping, primer extension and nucleotide sequence data also confirmed the exon / intron structure in mouse and hamster (Basler *et al.*, 1986; Westaway *et al.*, 1987).

After the failure of the cosmid library (see chapter 3) to yield a positive clone, a PrP clone, derived from a *Sinc* p7 mouse, became available from D. Westaway (see Westaway *et al.*, 1991 and figure 13). From the accompanying BamHI restriction map came the first indication that the PrP gene had two 5' untranslated exons. As a preliminary to identifying regulatory sequences within the PrP gene it was necessary to confirm this exon structure.

The main aims of this chapter are basically three-fold: i) To isolate a subclone, from the cosmid clone, which contains both 5' exons and the putative promoter region. ii) To determine a detailed restriction map of such a subclone. iii) To analyse the 5' flanking region and confirm the exon / intron borders by nucleotide sequencing.

The cosmid clone, itself, contains an insert of 40.8 kb and is flanked by NotI restriction sites. The insert is cloned into a cosmid vector called pcos6.EMBL (Ehrich *et al.*, 1987). The clone encompasses approximately 6.5 kb of 5' and 18 kb of 3' flanking sequences. Within these flanking sequences all three PrP exons are present. Exon 3 contains the entire open reading frame and is found within a 8.8 kb BamHI fragment (Westaway *et al.*, 1991).

The cosmid clone of the mouse PrP gene and the subclones made from it were used throughout this thesis. The data presented subsequently, in this chapter, was at the time novel and had not been shown previously.






Vector

Exon

K - KpnI (not all sites shown)

B - BamHI

**Figure 13.**

Organisation and structure of murine PrP cosmid clone.
Exons are numbered 1, 2 and 3. Exons 1 and 2 are 5' untranslated exons.
Exon 3 contains the entire open reading frame. Exon 3 also contains
10 bp of 5' untranslated region and approximately 1.1 kb of
3' untranslated region. Distances between BamHI sites are given in kb.

## 4.2 Isolation of subclones from the 5' region

To analyse the 5' region of the PrP gene it was necessary to obtain subclones from the cosmid clone.

The vector used for the initial subcloning was pT7T3 18U (Pharmacia). This vector is a phagemid and hence single stranded DNA can be produced by infecting a transformed culture with the helper phage M13 K07. The multiple cloning site of pT7T3 18 U is flanked by the T3 and T7 promoters. The nomenclature of the following subclones was based firstly on the vector (pT18 = pT7T3 18 U), then on the insert (e.g PrP gene fragments) and then finally on the size of the insert (e.g 3.5 = 3.5 kb).

a) pT18 PrP 9 b pT18 PrP 3.5 (subclone of pT18 PrP 9)

a) i) pT18 PrP 9

This plasmid comprises a 9 kb KpnI fragment from the cosmid clone which is cloned into the KpnI site of pT7T3 18U . The 9 kb KpnI fragment stretches from the KpnI site in the middle of exon 2 of the PrP gene to a KpnI site in the vector (see figure 13). Locht *et al.*, (1986) had previously shown that a KpnI site was located in the 5' non-coding leader sequence of a mouse cDNA clone. At this stage the existence of 5' non-coding exons in the mouse PrP gene was not known. However, it now seems very likely that this KpnI site corresponds to the KpnI site in exon 2. This was confirmed by digesting the cosmid clone with KpnI and using the 5.3 kb BamHI fragment, at the extreme 5' end of the insert (see figure 13), to probe a Southern blot. This probe hybridised only to a 9 kb restriction fragment. In addition, apart from the 5' exons the only other non-coding region upstream of the open reading frame in exon 3 is a stretch of 10 bp which directly proceeds the open reading frame. This region does not contain a KpnI site (Westaway *et al.*, 1987). Therefore this KpnI site must be in one of the upstream exons.

The 9 kb KpnI fragment was cloned into the KpnI site of pT7T3 18U in both orientations. These clones are called pT18 PrP 9a and pT18 PrP 9b.

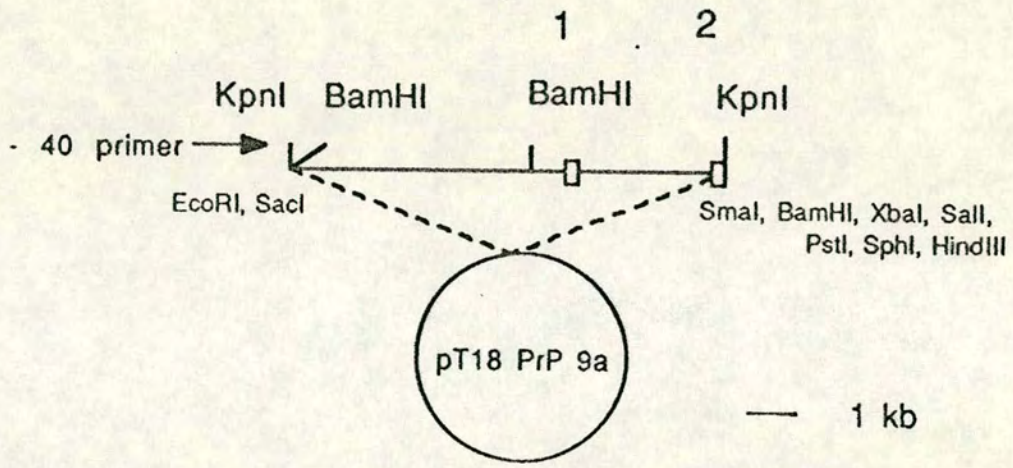



Figure 14.

Partial restriction map and orientation of 9 kb insert in pT18 PrP 9a. pT18 PrP 9b is in the opposite orientation than shown above. PrP exons are designated by open squares and are numbered 1 and 2. Some restriction sites are shown, as is the position of the target sequence to which the - 40 primer binds.

### 4.2 a) ii) Nucleotide sequence of region upstream of KpnI site in exon 2

Single stranded DNA was prepared from pT18 PrP 9b (see figure 14). This was sequenced using the -40 universal primer (5'-GTTTTCCCAGTCACGAG-3') as described in section 2.4.16. The sequencing was carried out to prove that exon 2 was present in the subclone. This could be done by determining its homology to mouse PrP cDNA (Locht *et al.*, 1986).

All sequence analysis described in this thesis was carried out on a VAX785 computer and the University of Wisconsin genetics computer group (uwgcg) sequence analysis software package was used throughout (Devereux *et al.*, 1984).

```
  1  gctgccctat ttcatgttca aatacgctcc atttgcgtga ctctttagta
 51  ttggtttgat gatttgcata ttagattaga ttgtatttca gttgtcagct
101  tatttatcaa ttctagtttt ctctttttgt tgttttaaag gactcctggt
151  atatttcaga actgaaccat ttcaaccgag ctgaagcatt ctgccttcct
201  agtgatcccg ggtaccc
```

**Figure 15.**

Nucleotide sequence of the region upstream of the KpnI site in exon 2 of the mouse PrP gene. Nucleotides numbered 1-159 represent intron sequence while nucleotides 160-217 represent PrP exon 2 sequence. This is based on homology to splice site consensus signals and to PrP cDNA which are detailed later in this chapter.

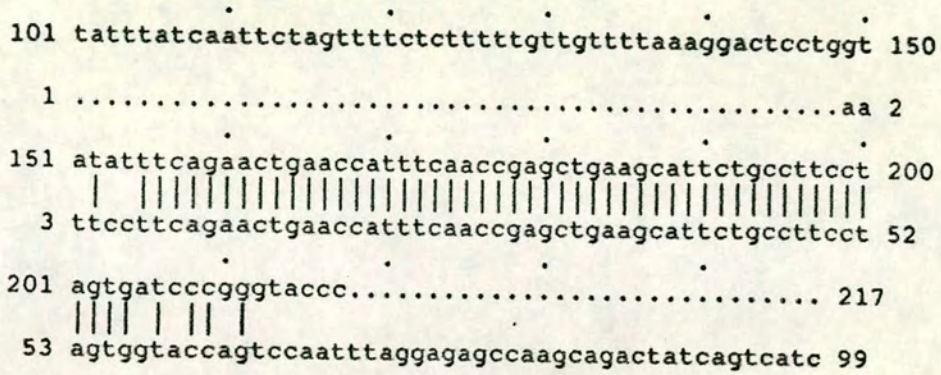



Figure 16.

Homology of region upstream of the KpnI site in exon 2 to mouse PrP cDNA. There is a 85 % homology between the sequence shown in figure 14 and the published mouse PrP cDNA (Locht *et al.*, 1986). PrP cDNA sequence is shown on the bottom lines and PrP intron 1 and exon 2 sequence is shown on the upper lines. The vertical lines between the two sequences represent areas of homology and gaps represent areas of mismatch. There is 100 % homology between nucleotides 155-204 of the genomic sequence (figure 14) and the corresponding region of the cDNA sequence. This is assumed to be PrP exon 2. The cDNA sequence of Locht *et al.*., (1986) is not full-length and does not include exon 1 (Westaway *et al.*, 1994 c).

4.2 b i) pT18 PrP 3.5

Having shown, by sequence homology to mouse PrP cDNA, that pT18 PrP 9 contained exon 2 of the PrP gene, the opposite end of the insert should contain 5' flanking region (see fig. 14).

This study aimed to examine the regulatory elements of the PrP promoter and the DNA sequences which constitute these elements. This was to be done, initially, by nucleotide sequencing. This strategy could be facilitated by subcloning a smaller fragment rather than working directly on the 9 kb insert of pT18 PrP 9 (fig. 14).

A 3.5 kb BamHI fragment was isolated from pT18 PrP 9a and cloned into the BamHI site of pT7T3 18U to produce pT18 PrP 3.5a and b. This 3.5 kb fragment stretches from the BamHI directly upstream of exon 1 to the BamHI site in the polylinker of the vector.

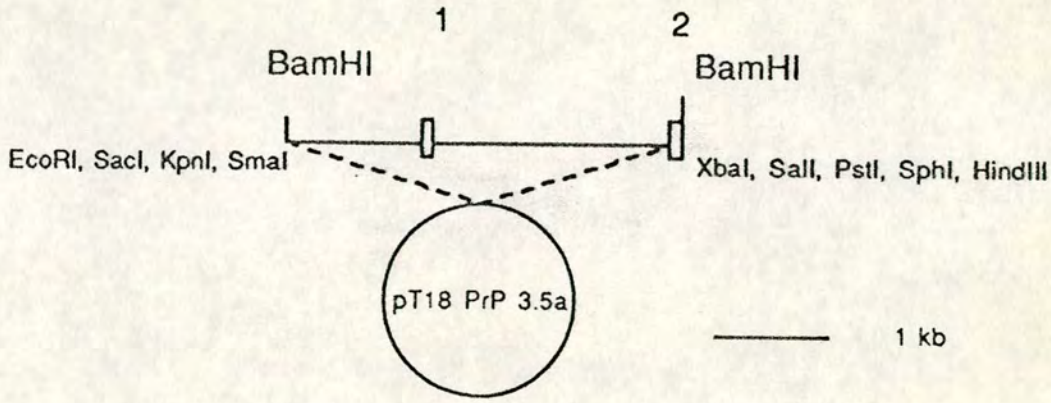



**Figure 17.**

Physical map of pT18 PrP 3.5 a. This clone contained approximately 1.2 kb of 5' flanking region, exon 1, and approximately half of exon 2. PrP exons are numbered 1 and 2. pT18 PrP 3.5 was made by cloning a 3.5 kb BamHI restriction fragment (from pT18 PrP 9) into the BamHI site of pT7T3 18U. The position of the restriction sites in the polylinker of the vector are given.

4.2 b ii) Restriction mapping of pT18 PrP 3.5

A restriction map was determined for pT18 PrP 3.5. This was done in order to aid in the sequencing of the 5' flanking region and in the subsequent deletion analysis of this region and of the first intron (see chapter 5).

Restriction sites were determined for EcoRI, HindIII, HincII and StuI. These enzymes were chosen as they were not expected to restrict the DNA too frequently. Also, these sites are present in the multiple cloning site of the vector and as a result may aid in any recloning process. Both single and double digests of all four enzymes were carried out and the sizes of the resultant bands determined (see section 2.4.3 and 2.4.4). Enzymes which only cut in the polylinker were also used in double digests e.g SmaI. Where possible (i.e in the 5' region) the restriction sites were confirmed by nucleotide sequencing (see section 4.2 b iii).

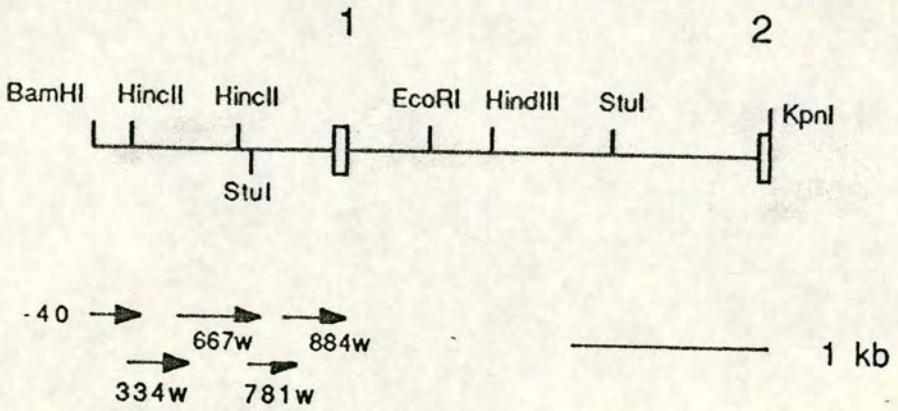



**Figure 18.**

Restriction map of pT18 PrP 3.5a. Position of restriction sites are shown. PrP exons are designated by open boxes and are numbered 1 and 2. In addition to the restriction map, the lower part of the figure shows the oligonucleotides which were used as primers for sequencing the 5' flanking region of the gene. The approximate length of sequence from each primer is indicated by arrows. See figure 19 for sequence of each oligonucleotide.

### 4.2 b iii) Nucleotide sequence of PrP 5' flanking region

By restriction mapping pT18 PrP 3.5 it was found that there were no suitable restriction sites which could be used to make internal deletions in the 5' flanking region (see figure 18). That is, there were no unique restriction sites (i.e which restrict the insert only once) separated from each other by ~ 200-300 bp in the 5' flanking region. If there were, one could have utilised them to create a construct which had deletions progressing from the 5' end of the flanking region to exon 1. Instead a series of oligonucleotides were synthesised and used as sequencing primers. For each primer, the sequenced samples were run on a denaturing 6 % polyacrylamide gel for 1.5 hrs (+ Mn buffer), 2.5 hrs (- Mn buffer) and 4.5 hrs (-Mn buffer). Mn buffer is used for obtaining sequence close to the primer. This ensured that maximal sequence data was obtained (for this apparatus) from one primer. This strategy was used to produce over 1.2 kb of sequence data from one strand of pT18 PrP 3.5. The sequence stretched from the BamHI site in the 5' flanking region to, and through, exon 1 and into the 5' end of the first intron.

This 1.2 kb of sequence obtained from the 5' flanking region of the mouse PrP gene (see figure 20) was analysed for binding sites of known transcription factors (figures 21 and 22), for the presence of a CpG island (figure 23) and also for its homology to PrP 5' flanking regions of other species (figures 24, 25, 26 and 27 ).

Synthetic oligonuleotides used as primers for sequencing PrP 5' region.

| | |
|---|---|
| - 40 | 5'- GTTTTCCCAGTCAGGAC-3' |
| 334w | 5'- TCAACTCAGGTCTACGC-3' |
| 667w | 5'- TGGGTGCGAGTCTACA-3' |
| 781w | 5'- ACGTGGCTTCCTCGCT-3' |
| 884w | 5'- CGCCATTGGTGAGCAT-3' |

**Figure 19.**

Sequence of synthetic oligonucleotides which were used as primers for sequencing the 5' flanking region of the PrP gene. Figure 18 shows the position of these oligonucleotides in relation to the 5' flanking region.

```
   1  TAGGCTAATA CTGGATACAA ATATTAATCC AAACCCAATC TTGTGTCTGT
  51  TAATGATCTT CAGTGTCTCG CCCTCAGCAA GAGGACAGGA TATTATGTTT
 101  TCCCTGTGAT TTATGACCTC TTCTGTCTCA GTATTGGCAG CAATTTATTT
 151  ACATGGCTTT GGAGTGTGTT ATATGTGTAG TATGGACATG AGGGTGCATG
 201  TCAACCTATG TGTGGAGGCC AGAGGTCAAT GTCATGTCTT GCCAATCACT
 251  GCCAGTGTCC TGGATCAACT CAGGTCTCAC GCTGGATGGA CTGAGCCAGT
 301  GCCCCAGCTC CTAACCCTCC CCTGTTTTAA AAAGGTCTCA TTATGTTGCC
 351  CAGGTCAGCC TTGAACTTGA GAGTCTCCTG ACTGCAGGCC TCTCATCCAA
 401  GTCAGCAGGC ATCTTGAACA AGAACATCAT TTCCTTTAAG CTGTTTCAGG
 451  CTGTGTTTGG TGGGAGCTGT TAAATGCAGT GCATTTTTCC TTTGGACACA
 501  ATAAAAAGAA AAAAGTGATT AATGAGTTGG GTGTGGTGGG TGCGAGTCTA
 551  CAATCTAGAA CTCAGGAGAT TGAGGGAGAA GCTTGCTCTG AGTTTGAGGT
 601  CAGCTTAAAT TACTTAGTAG GAACACCAGG CCAAATTGGG CTATGGGATT
 651  GTCTCCAAAG ATAAAGAAAA AAGGGAAGGA GAGAAAAGAA AAGAAAAAGA
 701  AAGGAAAGAA GGGGAAAAGA AGGAATCAGC AGAGAATAAT AAGTCAACAT
 751  GCAATGGCCA ATATACTTTC TAGGCCTCTA ATTCTTTTAT AGTTTGTGAG
 801  GGAAAATGTC GAAAATCTTC GTTACCAATT TCTTGTTACC AAAGTTCAAC
 851  GTGGCTTCCT CGCTCCGTTA GTAATTTCAT TTTCTCAACT ACCCATTATG
 901  TAACGGGAGC ATTGGGTACT GGATCAGTCT TCCATTAAAG ATGATTTTTT
 951  ATAGTTGCTG AGCGTCGTCA GGGAGTGCTG ACACTGGGGG CGGTTTAAAC
1001  AGATACAAGC ATTTAAGCCA GTCCGGAGCG GTGACTCACC CCCCCCCACC
1051  CCCACCCCCC CGCGAGAGAC GCGCGCGCCA TTGGTGAGCA TCACGCCCCG
1101  CCCCTCGCCT CCCGCCTAGC TCCCGCCTGC CCGCCCGCCC CTTTCCACTC
1151  CCGGCCCCCC CCGCGTTGTC GGATCAGCAG ACCGATTCTG GGCGCTGCGT
1201  CCGATCGGTG GCAG/GTAAGC GGGCTGCTGA AGCAGGCCTG GCGTTGCACT
1251  CAGCTTCGTG GTAAG
```

**Figure 20.**

Nucleotide sequence of the 5' flanking region of the PrP gene. Nucleotide number 1 indicates the region of DNA closest to the BamHI site directly upstream of exon 1(see figure 13). Potential 3' splice site is indicated with a slash and underlined and sequence 3' to this is intron.

## 4.3 Transcription factor binding sites in the 5' flanking region of the PrP gene

The 1.2 kb of 5' flanking sequence from the mouse PrP gene was analysed for the presence of potential transcription factor binding sites. Over 2000 transcription factor binding sites were used in the search. [database: TFsites1.dat (Ghosh , 1990)]. Over 140 sites for transcription factors were found within the 1.2 kb of 5' flanking region. Some of these sites are shown in figures 21 and 22. The potential binding sites are represented as vertical lines. The horizontal lines represent the 5' flanking sequence and nucleotides are numbered. Nucleotide number 1 represents the base nearest the BamHI site which is 1.2 kb 5' to exon 1. Nucleotide 1100 represents, approximately, the beginning of exon 1 (see chapter 6). The name of the transcription factor is given on the left hand side of each figure and the putative binding sequence is given on the right hand side. The number of times a binding site is found is also given on the right hand of each figure. Incompletely specified bases are shown in accordance with the nomenclature committee for the international union of biochemistry, 1985, (see abbreviations).

Within the 1.2 kb of PrP 5' flanking sequence there are potential sites for transcription factors such as Sp1 (Dynan and Tjian, 1983), AP1 (Angel *et al.*, 1987), AP2 (Mitchell *et al.*, 1987; Imagawa *et* al., 1987) and EGR-1 (Lemaire *et al.*, 1990; Cao *et al.*, 1990). There are approximately 140 potential binding sites for transcription factors within this 1.2 kb DNA sequence. However, which of these have a role in PrP gene expression is not known at this stage. From figures 21 and 22 it should be noted that the majority of transcription factor binding sites are located within, approximately, nucleotides 950-1150. These occur in the region where transcription initiation occurs and suggests that at least some of these transcription factors may play a role in the transcription of PrP (see chapter 6 and Westaway *et al.*, 1994 c).

It should also be noted that some transcription factors are represented more than once e.g Sp1. This is because Sp1 has a loose binding specificity (Faisst and Meyer, 1992) and as such this factor may recognise different sequences in different genes.

**Figure 21.** (see text for details)

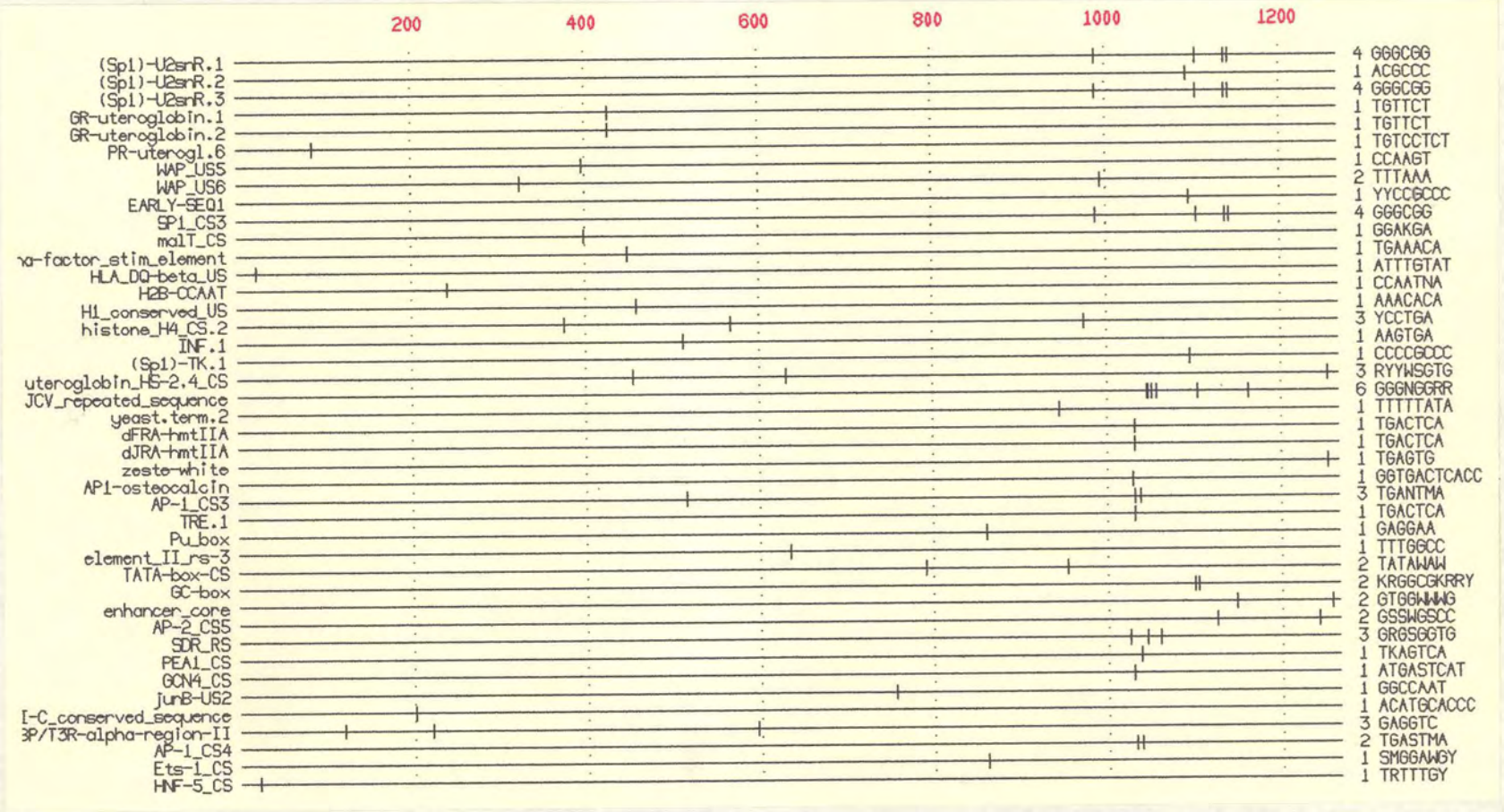
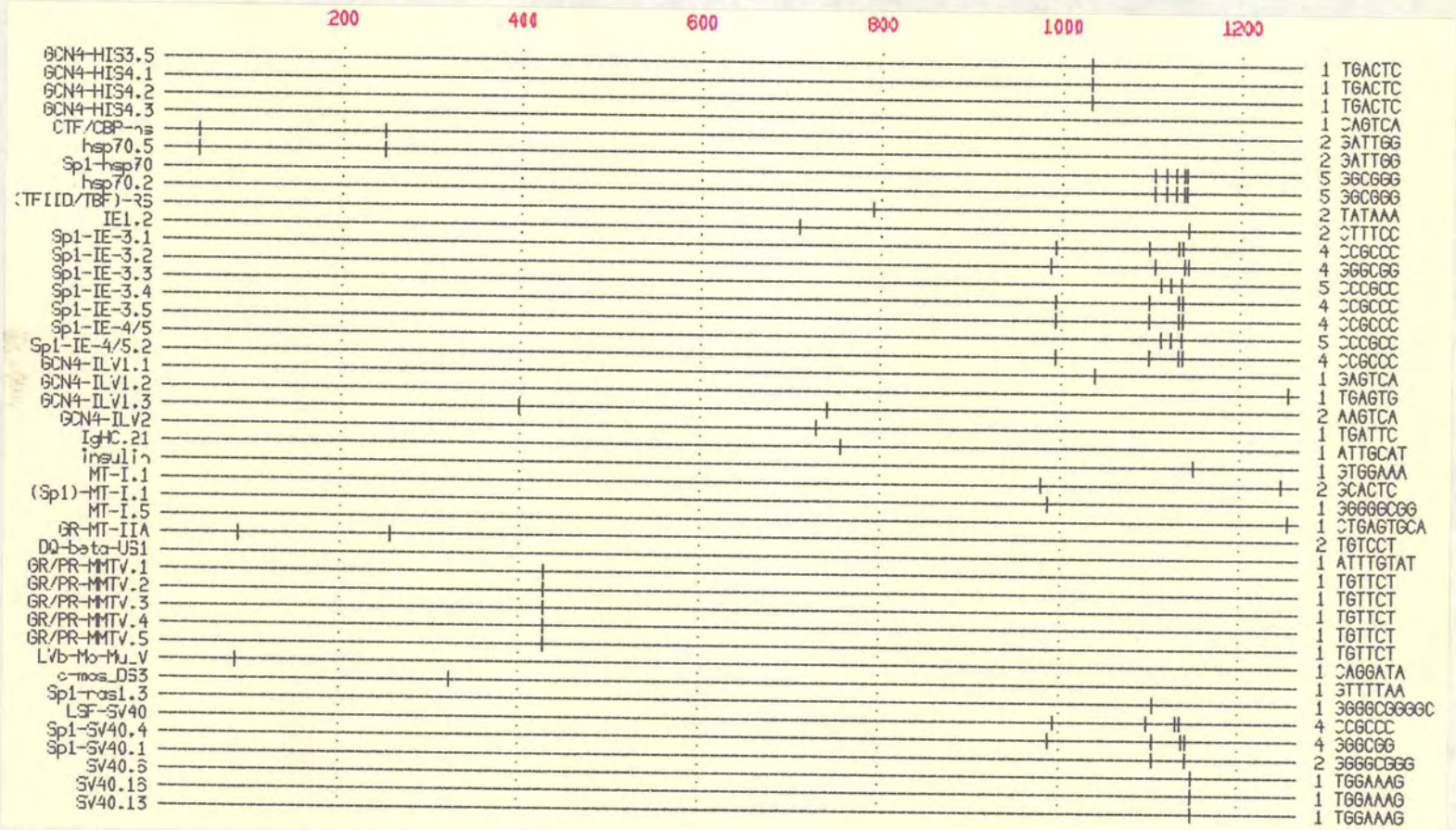


**Figure 22.** (see text for details)



## 4.4 CpG island in the 5' flanking region of mouse PrP gene.

In vertebrate DNA the frequency by which the dinucleotide CpG occurs is much lower than would be expected from the base composition (Josse *et al.*, 1961). The relatively rare occurrence of CpG dinucleotides is probably explained by the failure of cellular DNA repair mechanisms to recognise deamination of 5-methylcytosine to give thymine (Coulondre *et al.*, 1978). This is significant considering that between 60-90 % of all CpGs are methylated at the 5 position of the cytosine molecule (Bird, 1986). Despite the rarity of CpGs, there are regions of the genome where CpGs occur at close to the expected frequency. These regions are known as CpG islands and are predominantly unmethylated (Gardiner-Garden and Frommer, 1987).

The 5' flanking sequence (figure 20) obtained from the mouse PrP gene was analysed for the presence of CpG islands. This is the first such analysis for a PrP gene from any species. Although the methylation status of PrP 5' flanking region was not investigated, CpG islands are sufficiently distinctive to be identified by computer analysis (Gardiner-Garden and Frommer, 1978). The analysis involves plotting the observed / expected CpGs (obs / exp CpGs) against the average G+C percentage. The obs/exp CpG ratio can be calculated as shown below

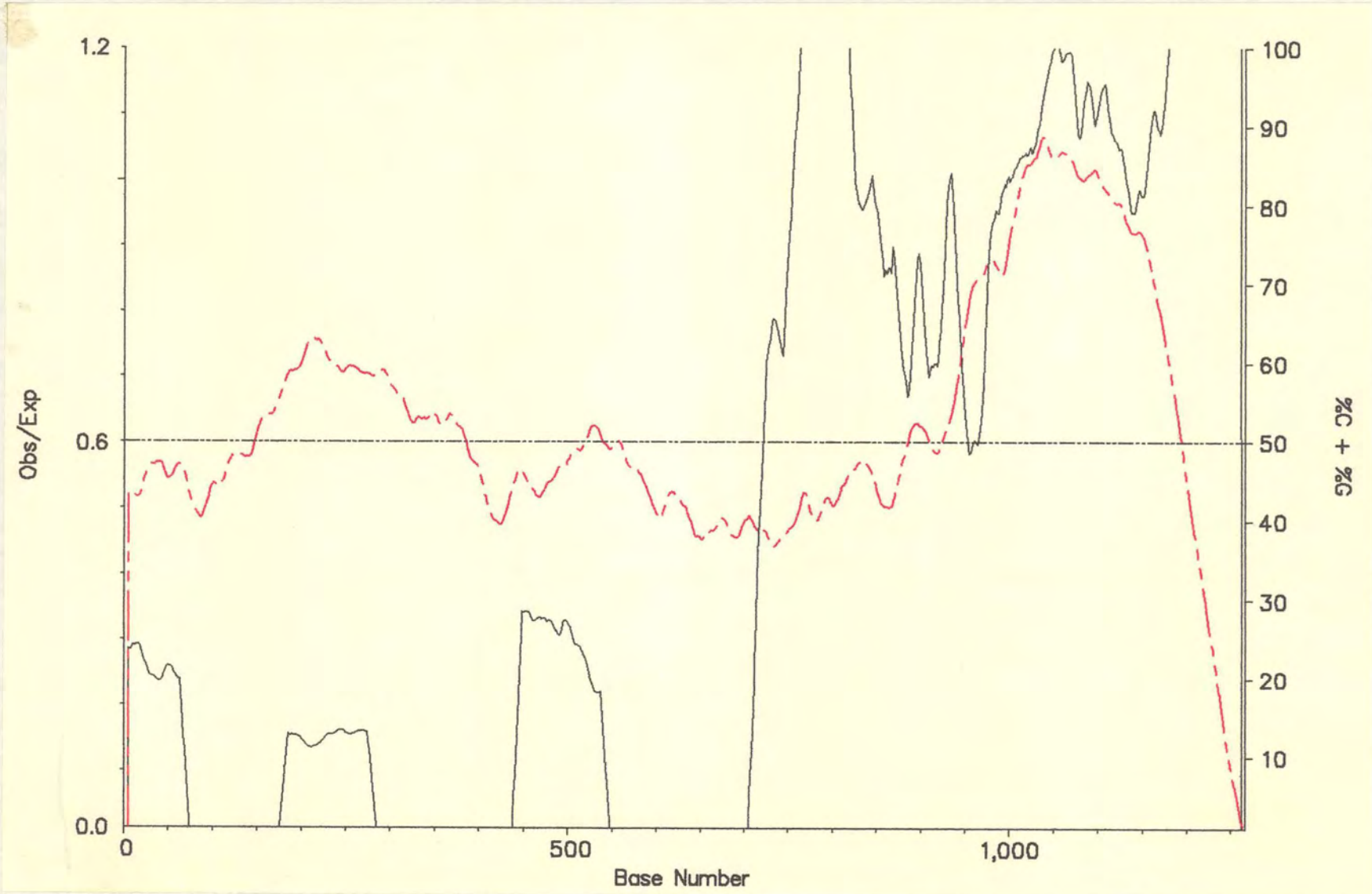
$$\text{Obs/Exp CpG ratio} = \frac{\text{Number of CpG}}{\text{Number of C x number of G}} \times N$$

where N is the total number of nucleotides. A moving average value for the percentage C+G (% G+C) and for the obs/exp CpG was calculated for 100 bp windows of sequence at 1 bp intervals.

A CpG rich region can be defined as a stretch of DNA with a moving % G+C over 50 and a obs/exp CpG ratio greater than 0.6. CpG rich regions which were at least 200 bp in length are classed as CpG islands, in accordance with Gardiner-Garden and Frommer (1978).

Figure 23.

CpG island occurrence in mouse PrP 5' flanking region. By using the 1.2 kb of 5' flanking sequence, a graph of the number of obs/exp CpGs was plotted against the % G+C. The broken line represents the % G+C and the solid line represents the obs/exp CpGs. Based on criteria of Gardiner-Garden and Frommer (1978), which is described in the text, a CpG island is located between nucleotides 970-1180. This CpG island extends from the 5' non transcribed flanking region to exon 1.


1.2
0.6
0.0
Obs/Exp
0
500
1,000
Base Number
100
90
80
70
60
50
40
30
20
10
%C + %G

The length of the CpG island found in the 5' flanking region of the murine PrP gene (and exon 1) is at least 210 bp in length, based on the sequence data which were presented in figure 20 of this chapter. It may be that the CpG island is larger than this and extends well into the gene, as is the case with most 5' CpG islands (Gardiner-Garden and Frommer, 1978). However, more sequence information from the first intron of the PrP gene would be needed to prove this. Indeed, in this same study they found that there were no examples whereby a CpG island lay entirely within the 5' untranslated flank of a gene.

## 4.5 Homology of mouse PrP 5' flanking region to PrP 5' flanking regions of other species

**a) mouse to human**

```
951 ATAGTTGCTGAGCGTCGTCAGGGAGTGCTGACACTGGGGGCGGTTTAAAC 1000
      || |  || |  | | | | ||  | | || | | ||  |
  1 .....ggccggccgcccgccgcgggggcacagagtgtgcgccg....... 38

1001 AGATACAAGCATTTAAGCCAGTCCGGAGCGGTGACTCACCCCCCCCCACC 1050
                           ||   ||| || |    |||| || |
 39 ......................gcgcgccggcgaattggtccccgccgcg 66

1051 CCCACCCCCCCCGCGAGAGACGCGCGCGCCATTGGTGAGCATCACGCCCCG 1100
      || | | ||||||||  | ||| ||  || ||         | ||||||||
 67 acctc.cgcccgcgagcgccgcccgttcccctt........ccccgccccg 107

1101 CCCCTCGCCTCCCGCCTAGCTCCCGCCTGCCCGCCCGCCCCTTTCCACTC 1150
     |  | | ||  | |  |   | | ||| |  || || | | | | ||| ||
 108 cgtccctccccgcctcggccccgcgcgtcgcctccagtcgc.tgccagtc 156

1151 CCGGCCCCCCCCCGCGTTGTCGGATCAGCAGACCGATTCTGGGCGCTGCGT 1200
      | | |  || ||             ||    |||   |   | || |
 157 gctgacagccgcg.............gcgccgcgagcttctcctcctcctc 193

1201 CCGATCGGTGGCAGGTAAGCGGGCTGCTGAAGCAGGCCTGGCGTTGCACT 1250
      |||  | | |||||||||| ||  | |
 194 acga.ccgaggcaggtaaacgcccgggg..................... 220
```

**Figure 24.**

Homology between human 5' flanking region (Puckett *et al.*, 1991) and mouse 5' flanking region (this study). Mouse sequence is shown in uppercase while human sequence is shown in lowercase. These sequences show a 58 % homology to each other (using the gap programme: gap length = 3, length weight = 0.1). Regions of identity are shown as vertical lines between nucleotides.

b) mouse to hamster

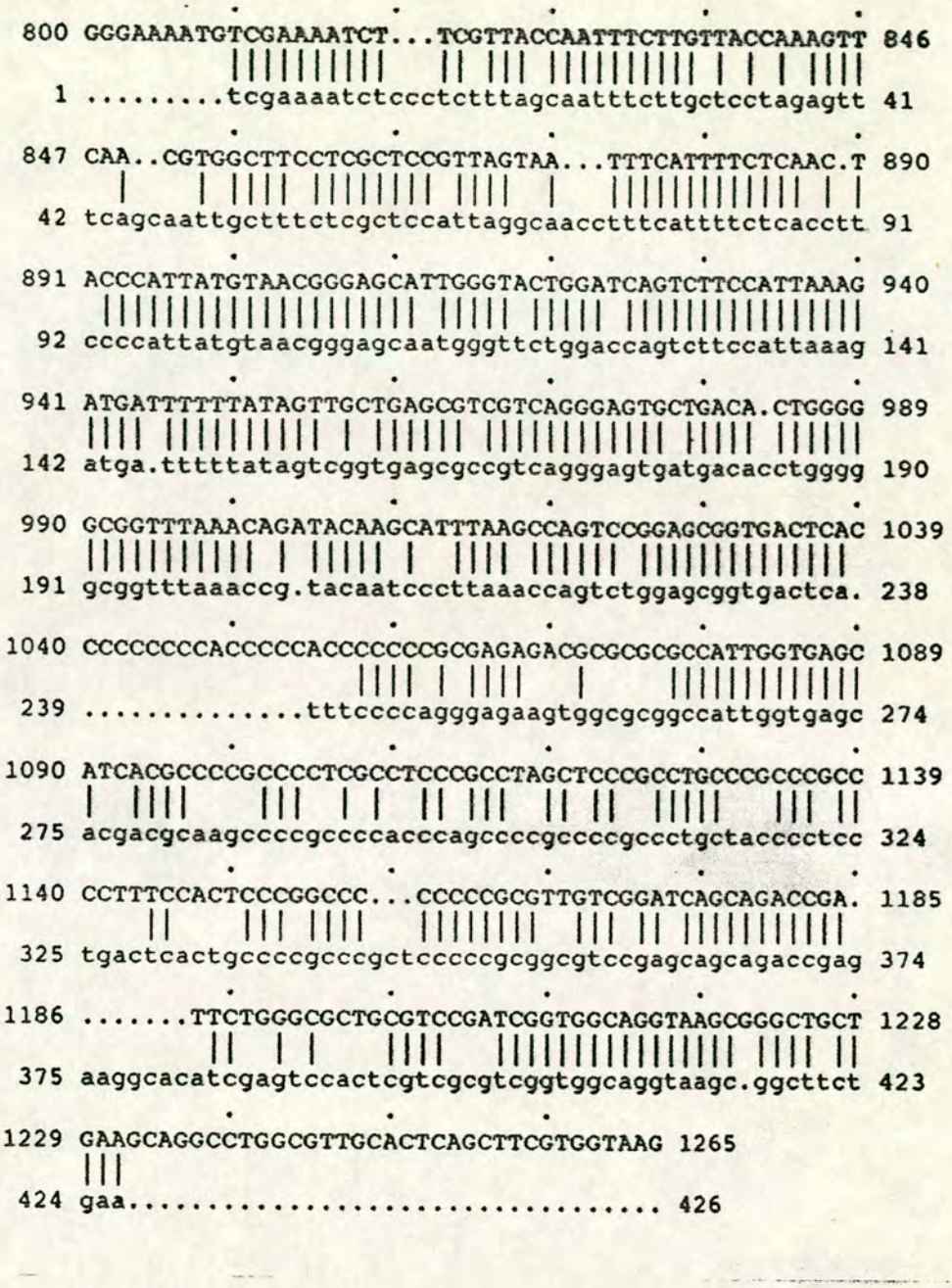



**Figure 25.**

Homology between hamster 5' flanking region (Basler *et al.*, 1986) and mouse 5' flanking region (this study). Mouse sequence is shown in uppercase while hamster sequence is shown in lowercase. These sequences show a 78 % homology to each other (using the gap programme: gap length = 3, length weight = 0.1). Regions of identity are shown as vertical lines between nucleotides.

c) mouse to sheep

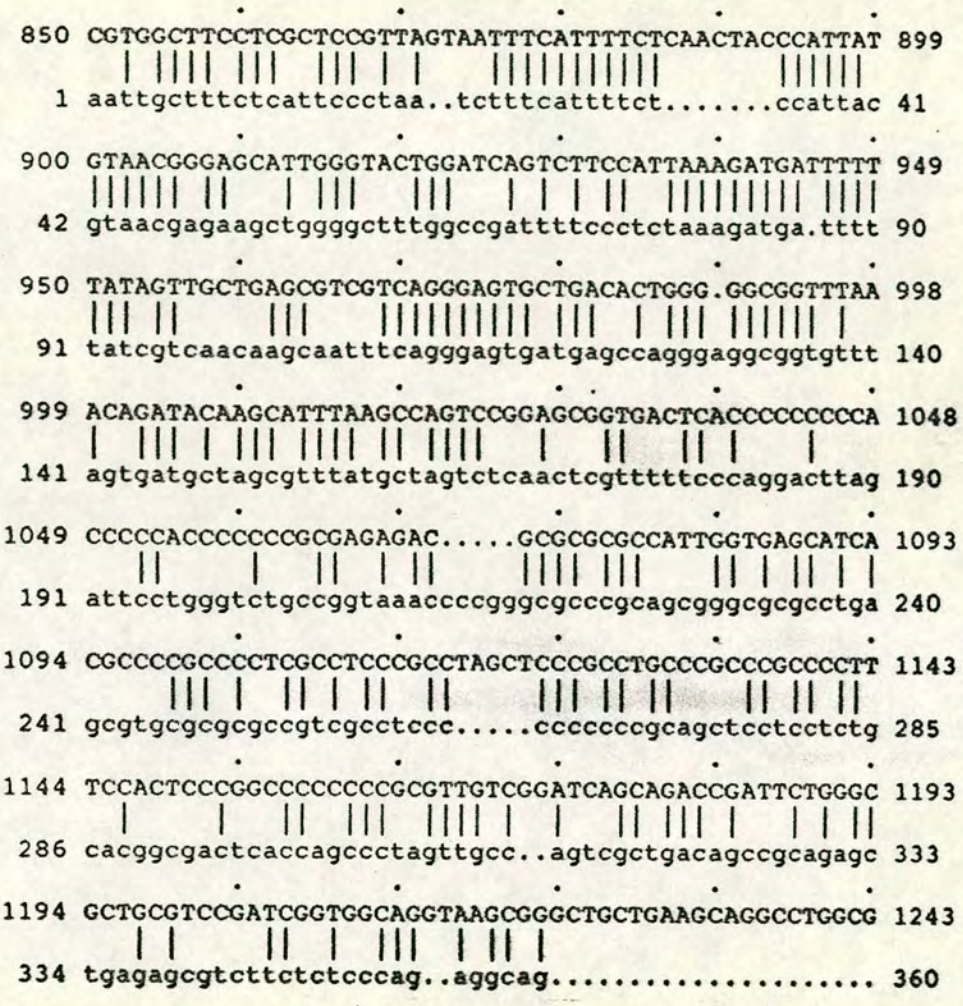



**Figure 26**

Homology between sheep 5' flanking region (Westaway *et al.*, 1994 b) and mouse 5' flanking region (this study). Mouse sequence is shown in uppercase while human sequence is shown in lowercase. These sequences show a 57 % homology to each other (using the gap programme: gap length = 3, length weight = 0.1). Regions of identity are shown as vertical lines between nucleotides.

d) mouse (published) to mouse (this study)

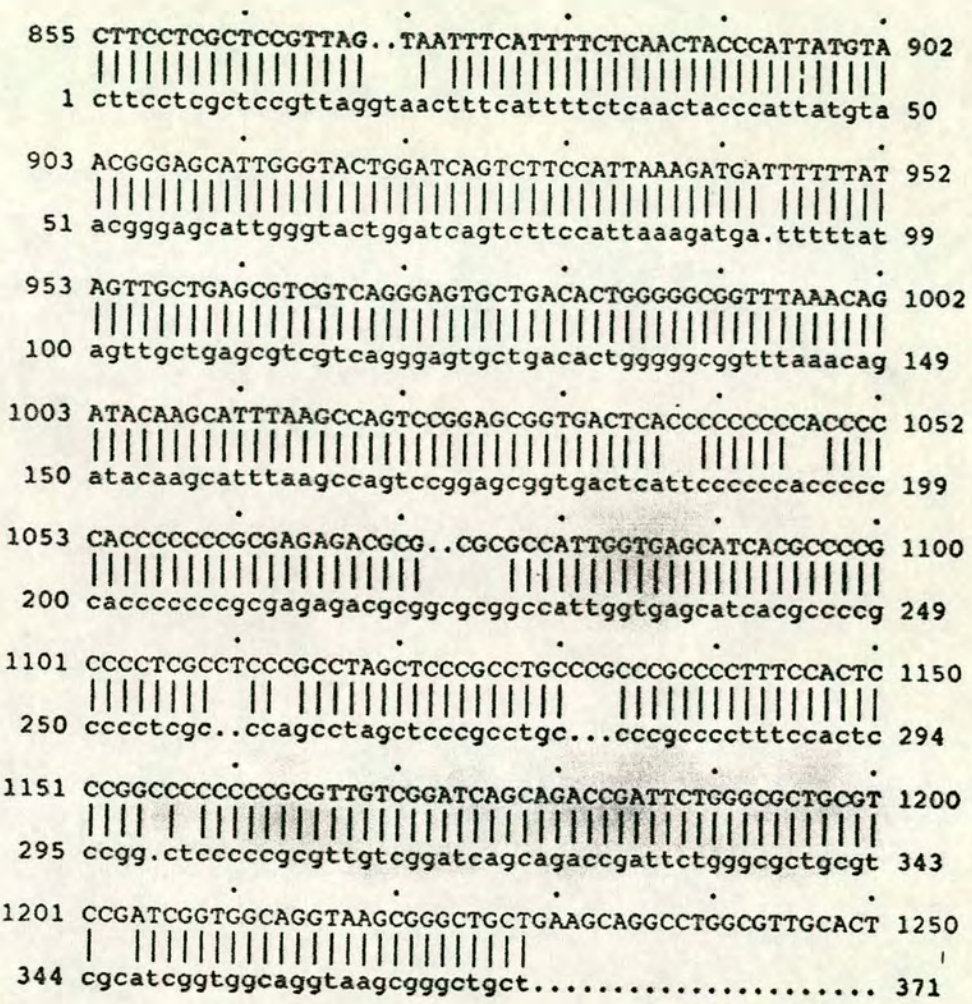



Figure 27

Homology between mouse 5' flanking region (Westaway *et al.*, 1994 c) and mouse 5' flanking region (this study). Mouse sequence produced in this study is shown in uppercase while the published sequence is shown in lowercase. These sequences show a 96.5 % homology to each other (using the gap programme: gap length = 3, length weight = 0.1). Regions of identity are shown as vertical lines between nucleotides.

The 5' flanking regions of mouse and hamster PrP (Basler *et al.*, 1986) genes are well conserved, showing a homology of 78 % (figure 25) over a 400 bp region of 5' flanking region. This level of homology between two closely related species is not unexpected and may represent similar regulatory processes between the two species. Westaway *et al.*, (1994 c) have recently published the PrP 5' flanking sequence from a cosmid clone derived from a I/LnJ mouse. In the present study, subclones from the same cosmid clone were also sequenced. The sequence produced in this study is 96.5 % homologous to the sequence produced by Westaway *et al.*, (1994 c). Since both sequences were generated from the same strain of mouse and indeed from the same cosmid clone, one may have anticipated a higher level of identity between the two sequences. In this chapter the strategy employed to sequence the PrP 5' flanking region has been described in detail: briefly, the 5' flanking region was sequenced on one strand using five oligonucleotides to generate 1.26 kb of sequence. The sequencing strategy employed by Westaway *et al.*, is, however, not given in detail and one cannot ascertain if one or both strands were sequenced. We do know, however, that double stranded DNA was used as the template for sequencing in that study. This form of sequencing is widely known to be less reproducible than single stranded sequencing; producing more variable band intensities and less readable sequence further from the primer (Sambrook *et al.*, 1989). Such differences in sequencing strategy may account for the differences when the two sequences are compared (figure 27). This figure shows that most differences are minor, representing one or two nucleotide mismatches. There are, however, larger differences between the two sequences. For example, there is a mismatch of five nucleotides between nucleotide numbers 1075-1079 of my sequence and numbers 222-226 of the Westaway *et al.*, sequence (figure 27). On sequencing autoradiographs there appears to be no ambiguities in this area, as shown in figure 28.

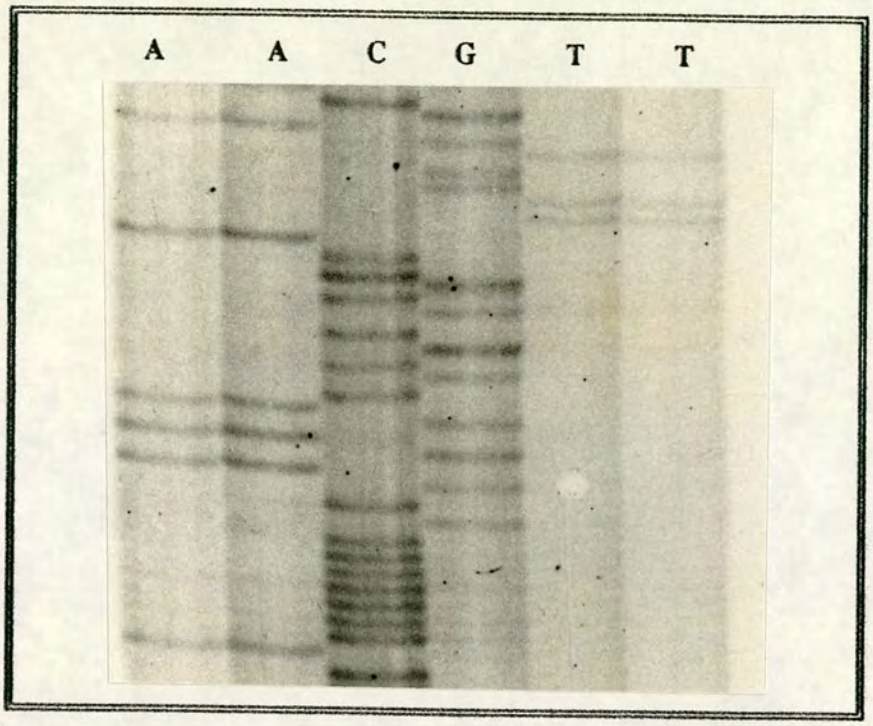



**Figure 28.**

Autoradiograph of sequence derived from PrP 5' flanking region between nucleotides 1054-1089. Sequence was generated from single stranded DNA, derived from pT18 PrP 3.5, using oligonucleotide 781 ( 5' ACGTGGCTTCCTCGCT 3' ). Nucleotides CGCGC corresponding to nucleotides numbers 1075-1079 are clearly present, but are absent from the sequence of Westaway *et al.*,. (1994 c), as illustrated in figure 27.

The discrepancies between the two sequences, as depicted in figure 27 and 28, were probably due to inaccuracies in the reading of sequencing autoradiographs, especially at highly G C regions. These discrepancies may have arisen because of errors in either my work, Westaway's work, or both.

From figures 26 and 27, one can see that homology between mouse and sheep and human PrP 5' flanking regions is more limited. Mouse and sheep are 57 % homologous whereas mouse and human are 58 % homologous.

### 4.6 Conservation of splice sites

Potential spice sites have been determined for the 5' and the 3' splice sites of the first intron of the PrP gene. These are based on homology to the consensus sequences derived by Mount (1982).

As mentioned previously, at the time this work was initiated there was no direct evidence that the mouse PrP gene had two 5' exons. However, through comparing sequence data in exon 1 and 2 with splice site consensus sequences this shows that this is the case and the recent study by Westaway *et al.*, (1994 c) confirms this.

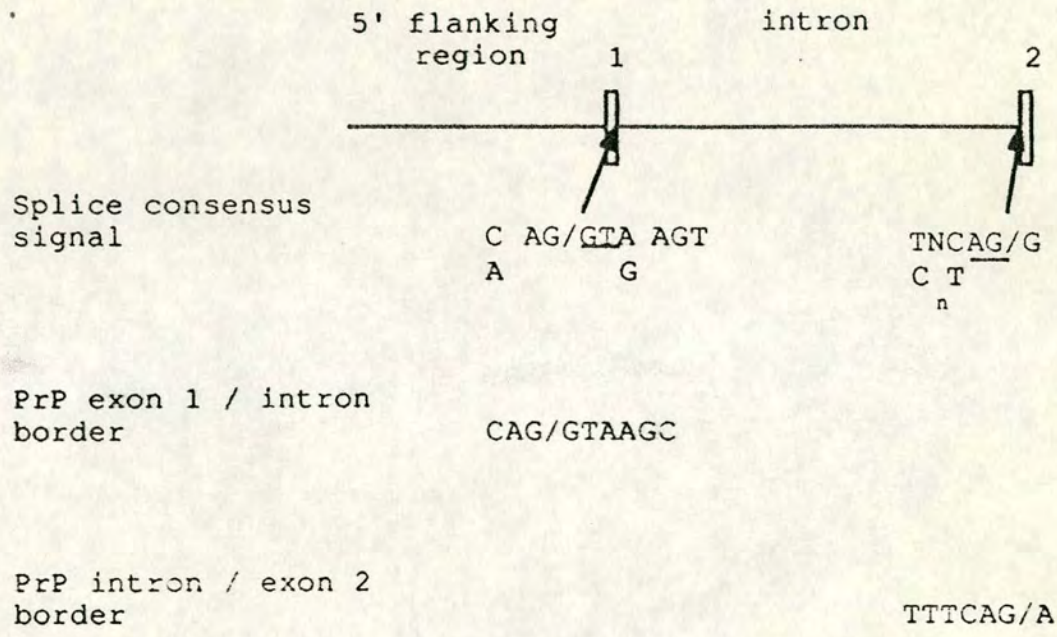



**Figure 29.**

Homology of PrP splice sites to donor and acceptor splice site consensus sequences (Mount, 1987). The underlined nucleotides of the splicing consensus signal are the first two (GT) and the last two (AG) nucleotides of an intron and these are strictly conserved. The PrP splice sites are in good agreement with consensus splice sites, both of which show one mismatch. Exons are numbered and are represented as open boxes.

## 4.7 Determination of the 1st intron length

The length of the BamHI insert of pT18 PrP 3.5 was measured at 3.5 kb. This was sequenced from the BamHI in the 5' flanking region through to exon 1 and into the intron. There were 1214 nucleotides from the BamHI site to the exon 1 / intron border. This border was identified by sequence homology; a splice site was found at nucleotide number 1214 (CAG/GTAAGC). Similarly, a splice site was found at the intron / exon 2 border at nucleotide number 140 of figure 11 (TTTCAG/A). This showed that there were approximately 60 nucleotides of exon 2 present in pT18 PrP 3.5

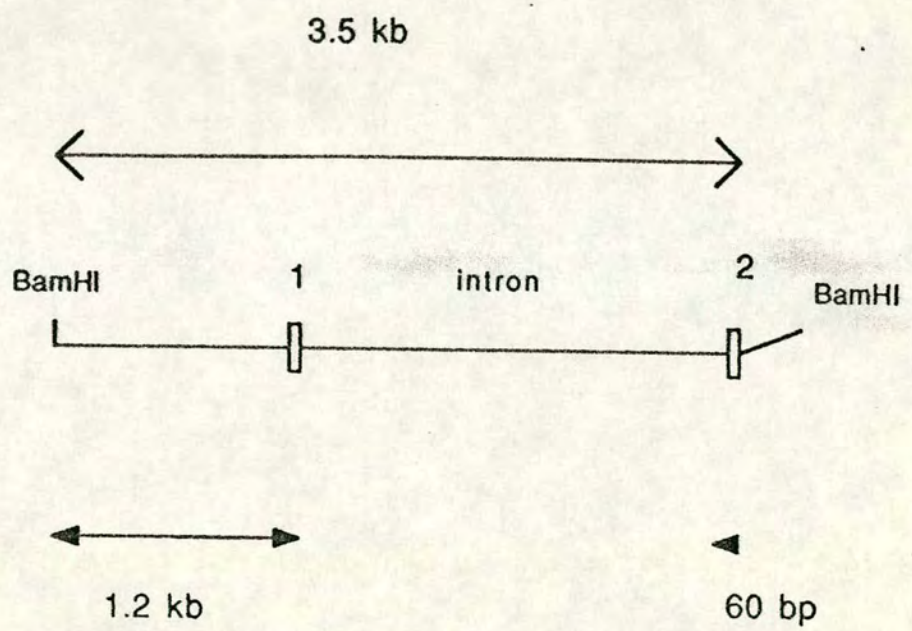



**Figure 30.**

Determination of intron 1 length based on sequence homology to splice consensus signals. The entire insert of pT18 PrP 3.5, derived from the 5' end of the PrP gene, is approximately 3.5 kb in length. From the 5' BamHI site to the exon 1 / intron border is 1214 bp. 60 bp of exon 2 are also present within the insert. Then remainder is 2.2 kb and this is assumed to be intron 1.

## 4.8 Discussion

To recapitulate, briefly, a subclone designated pT18 PrP 3.5 was analysed in detail by nucleotide sequencing. This clone contains exon 1, intron 1, approximately 60 bp of exon 2 and 1.2 kb of 5' flanking sequence.

The entire 1.2 kb of 5' flanking region was sequenced on one strand using sequential oligonucleotides as primers. Analysis of this sequence revealed several features. Firstly, there is quite a large variation in the G+C richness over the 1.2 kb of sequence. For example, at the extreme 5' end (nucleotides 1-100) the G+C content is only 38 %. However, nearer exon 1 (nucleotides 1101-1219) the 5' flanking region has a G+C composition of 75 %. In addition, the sequence data shows that there is no apparent TATA box located 20-30 bp upstream of exon 1 (Breathnach and Chambon, 1981), nor is there a CAAT box which is usually present further upstream of the TATA box (Benoist and Chambon, 1980).

Computer analysis of the 5' flanking region showed that a CpG island exists between nucleotides 970-1180. This CpG island may extend further into the gene (i.e intron 1 and exon 2) but in the absence of more extensive sequence data, this is impossible to ascertain. The existence of a CpG island in the 5' flanking region of the PrP gene may provide a mechanism whereby the expression of PrP could be regulated. For example, it is known that PrP mRNA is expressed more than 100 times more highly in the brain than in the spleen (Caughey *et al.*, 1988). Although the mechanisms controlling this differential expression have not been investigated, it may that methylation levels of the CpG island in different tissues vary. The CpG island in the brain, for example, would be poorly methylated, allowing increased transcription levels. The CpG island of the spleen, on the other hand, would be extensively methylated, thus inhibiting transcription of PrP. When mouse genomic DNA from spleen and brain was restriction digested with MspI and HpaII, Southern blotted and probed with a 1.7 kb BamHI / EcoRI PrP fragment encompassing the 5' flanking region (see figure 18), differences could be detected in band size between brain and spleen. This data was not shown due to the poor

quality of the autoradiograph, but at least provides preliminary evidence that methylation differences exist in the PrP 5' flanking regions of spleen and brain.

Again by computer analysis, it was possible to identify potential binding sites for transcription factors within the PrP 5' flanking region. The potential binding sites for over a hundred different transcription factors were identified in this way.

The presence of such consensus sites does not, of course, prove that these transcription factors actually bind in such locations. To determine if proteins bind at such locations one could analyse the region by DNaseI footprinting, or at least initially by band shift assay.

Finally, by sequence analysis, it was shown that the 5' flanking region of the murine PrP gene highly homologous (78 %) to the hamster PrP gene 5' flanking region (Basler *et al.*, 1986), over a 400 bp stretch. This degree of conservation did not, however, extend to sheep (57 %) and human (58 %). Whether these differences in homology represent differences in the mechanisms by which PrP expression is controlled in different species, is not known. It may just represent evolutionary divergence. Such comparisons, though, could be more revealing if longer stretches of 5' flanking region were available for other species. Conserved sequences within these regions, across a number of species, would give a better indication of their functional importance. However, to prove functional importance of specific DNA sequences it is necessary to mutate them, for example by deletion. The effect of such mutations on transcription can then be measured.

The next chapter describes the construction of a number of deletions in the 5' flanking region and in the first intron and measures their effect in transient transfection assays.

# Chapter 5 Development of reporter gene constructs for the PrP gene and their analysis by transient transfection in neuro2a cells.

## 5.1 Introduction

Our knowledge of eukaryotic gene regulation has been greatly enhanced by the ability to introduce genes or fragments of genes into living systems: both transgenic mice and tissue culture cell lines.

The development of reporter gene vectors has proved invaluable in such studies. Reporter genes code for proteins which are exogenous or are easily distinguishable from all other cellular proteins found, for example, in a particular cell line. Reporter genes are most commonly used for the analysis of putative promoter regions, although they may also be used to study enhancers and repressors. The basic strategy for analysing promoter regions is to clone such DNA upstream of the coding sequence of the reporter gene. As a result, the reporter gene is placed under the transcriptional control of the test DNA. The influence of the test DNA can be quantitatively measured by introduction of the recombinant plasmid into a suitable cell line and assaying for the reporter gene product. The main advantage for using reporter constructs to study gene regulation is that the reporter protein is not found in the test system. This allows reporter gene expression to be distinguished from expression of the gene which is being studied. At present there are about five commonly employed reporter genes. These are chloramphenicol acetyl transferase (CAT), β-galactosidase, luciferase, alkaline phosphatase and human growth hormone (for review see Alam and Cook, 1990).

There are a number of different methods whereby recombinant reporter gene vectors may be introduced into mammalian tissue culture cells. These include calcium phosphate precipitation (Graham and Van der Eb, 1973; Chen and Okayama, 1987), lipofection (Felgner *et al.*, 1987), electroporation (Potter *et al.*, 1984) and DEAE mediated transfection (McCutchan and Pagano, 1968). There are also more specialised transfection methods such as microprojectile mediated bombardment (Klein *et al.*, 1988) and laser mediated transfection (Kurata *et al.*, 1986). These methods, however, are less commonly used.

This chapter aims to identify regulatory regions within the promoter and first intron of the PrP gene. This will be done by attaching various deleted fragments of the PrP gene to a reporter gene and then assaying for reporter activity after transient transfection of such constructs into neuro2a cells.

In this study, CAT was used as the reporter gene. CAT is an E.coli gene whose function is to confer resistance to the antibiotic chloramphenicol. It does this by transferring an acetyl group from acetyl co-enzyme A (coA) to chloramphenicol. CAT is not found in mammalian cells and it is easily assayed. The assay is linear over a large range of enzyme concentrations which means that it can be used for assaying promoters which have large variations in strength (Seed and Sheen, 1988). For these reasons CAT is the most widely used reporter gene.
For transfections, the calcium phosphate method was routinely used. This technique relies upon calcium chloride forming a precipitate in the presence of a phosphate buffer and DNA. This DNA / precipitate complex is presented to the cells and taken up by endocytosis. It is simple to perform and was found to be quite reproducible.

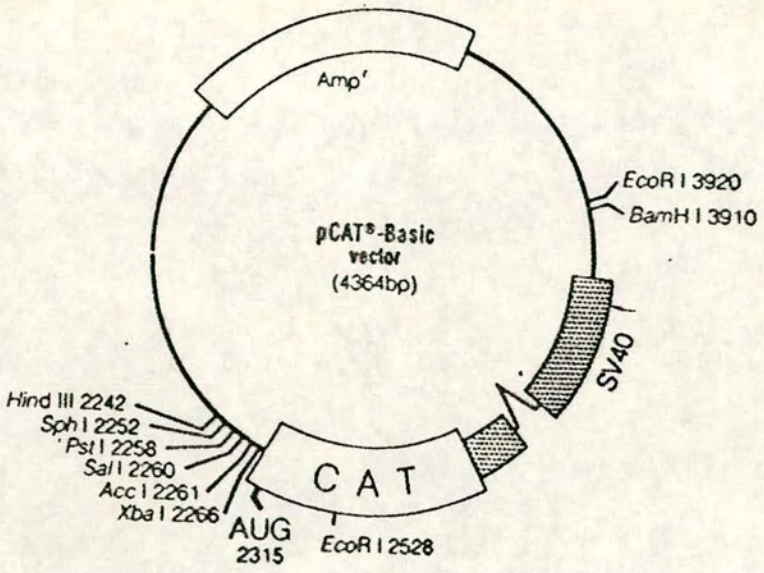



**Figure 31.** (figure taken from Promega)

pCAT basic (Promega) was the reporter vector used for PrP promoter analysis (Rosenthal, 1987). The vector is 4.364 kb in length and lacks both eukaryotic promoter and enhancer sequences. This allows putative promoter fragments to be cloned into the multiple cloning site or putative enhancer/repressor sequences to be cloned into either the multiple cloning site or the BamHI site which is downstream from the CAT gene. pCAT basic also has a gene for ampicillin resistance and a binding site for the reverse primer, located 5' to the multiple cloning site. The SV40 small T antigen is present 3' to the CAT coding region and contains an intron which enhances expression.

## 5.2 Construction of pCAT PrP 3.5

The entire 3.5 kb BamHI fragment from pT7T3 18 PrP 3.5 (section 4.2 b) was tested for promoter activity by firstly cloning into the XbaI site in the polylinker of pCAT basic. This was done by digesting pCAT basic with XbaI and then filling in the resulting 3' termini with klenow fragment to produce blunt ends (see section 2.4.8 a). The 3.5 kb BamHI fragment was excised from pT7T3 18 PrP 3.5 and was treated in the same fashion as above to produce a blunt ended fragment. Both were ligated together to form pCAT PrP 3.5.

The orientation of the insert was confirmed by restriction digests and by double stranded nucleotide sequencing using the reverse primer. The binding site for the reverse primer is located 15 bp upstream from the HindIII site in the polylinker of the vector. Depending on the orientation of the insert one would obtain either PrP exon 2 sequence (figure 15) or PrP 5' flanking sequence (figure 20) using this strategy. PrP 5' flanking sequence was shown to be present directly downstream of the polylinker. Therefore the insert was in the correct orientation with regard to promoting transcription of CAT.

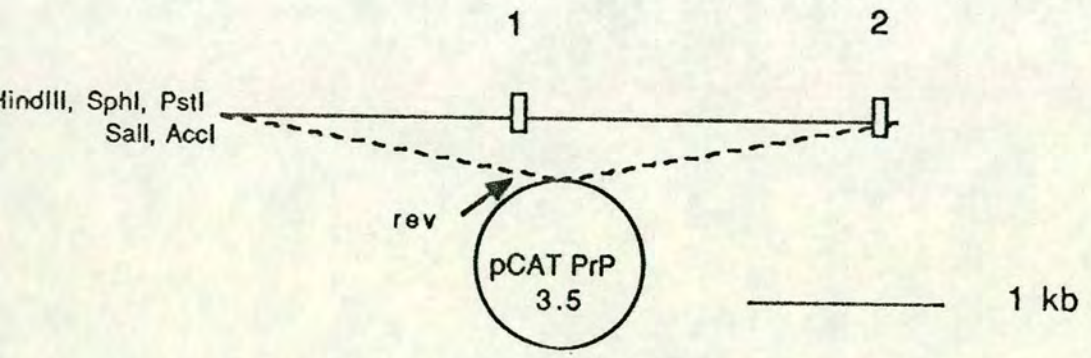



**Figure 32**

pCAT PrP 3.5 is a 7.8 kb construct which contains a 3.5 kb PrP fragment cloned upstream of the CAT gene in pCAT basic. This 3.5 kb PrP fragment contains 1.2 kb of 5' flanking region, exon 1, 2.3 kb intron, and about 60 bp of exon 2. Exons are numbered and the reverse primer binding site is indicated. The CAT gene is directly downstream from PrP exon 2.

## 5.3 Construction of deletion mutants

In order to define any regulatory elements which may exist within the promoter and intron, deletions were made in these regions. The rationale behind this is that if any regulatory elements were present within a deleted region then this would result in a different CAT value when compared to pCAT PrP 3.5.

Two deletions were made in the 5' flanking region and another was made in the intron, producing four different PrP CAT constructs. How these constructs were made is described below.

### a) Using exonuclease III

#### i) pCAT PrP 3.1

Initially, it was intended that all deletions in the 5' flanking region would be made using exonuclease III. This enzyme can digest DNA from a 5' protruding end or a blunt end, whereas 3' protruding ends are resistant to such digestion (Heinikoff, 1984). Deletion construction based on the exonuclease method has the advantage that the deletions can be very specific. This is because exonuclease III will digest DNA at a predictable rate when incubated at a given temperature.

Both SphI and SalI were shown to digest PCAT PrP 3.5 only once; in the multiple cloning site. SphI forms, exonuclease III resistant, 3' overhangs on digestion whereas SalI produces, exonuclease III susceptible, 5' overhangs. Therefore, when pCAT PrP 3.5 was digested with SalI and SphI, exonuclease can digest in a unidirectional manner into the insert. The rate of exonuclease digestion can be controlled by the incubation temperature. Generally, the lower the incubation temperature the slower the rate of exonuclease digestion. After exonuclease III digestion the DNA can be treated with S1 nuclease to remove single stranded tails left over from the exonuclease digestion. The digested DNA can then be made blunt, religated and transformed.

Protocol:

a) 10 µg of pCAT PrP 3.5 was cut to completion with SalI and SphI. The digested DNA was phenol:chloroform extracted, ethanol precipitated and washed in 70 % ethanol.

b) The DNA pellet was resuspended in 120 µl of 1 x exonuclease III buffer. 20 µl of this solution was prewarmed to 37°C and 5 µl (325 units) of exonuclease III added, mixed and returned to 37°C. At this temperature exonuclease III should digest at ~450 bp / min.

c) 2.5 µl samples were taken every 30 sec and mixed with 7.5 µl of a solution containing 9 µl S1 7.4 x S1 buffer, 20 units S1 nuclease, 57 µl $H_2O$.

d) After collection all of the samples then were incubated at room temperature for 30 mins.

e) 1 µl of S1 stop buffer was added and mixed to each time point aliquot. The S1 nuclease was heat inactivated at 70°C for 10 mins.

f) Digestion was checked on a 0.8% agarose gel by loading 1µl of each time point. The ends of the DNA were blunt ended, religated and transformed into E.coli CMK 603.

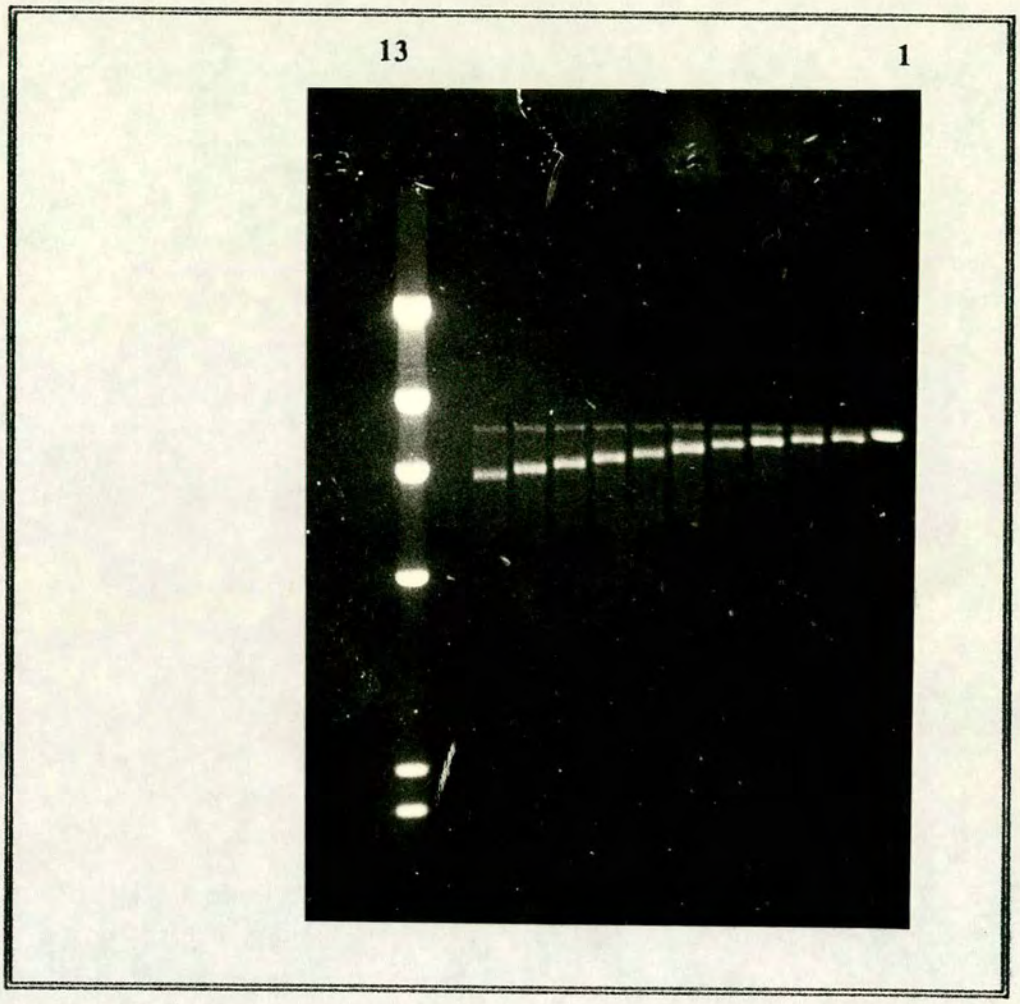



**Figure 33.**

Exonuclease III digestion of pCAT PrP 3.5.

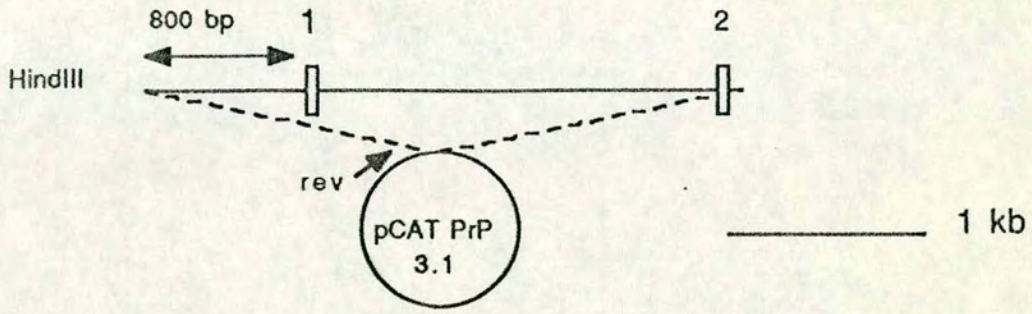
10 µg of pCAT PrP 3.5 was digested with SalI and SphI. 325 units of exonuclease III were added and incubated at 37°C for various time intervals. After heat inactivation at 70°C for 10 min the digest was incubated with 20 units of S1 nuclease at room temperature.

Lane 1-11: 200 ng of each time point (Lane 1 = 0 min, lane 2 = 1min etc)

Lane 13 : Lambda HindIII markers

pCAT PrP 3.1 was derived from the 1min time point when incubated with exonuclease III at 37°C.

A)




**B) TCAGCAGGCATCTTGAACA**

Partial nucleotide sequence data from the 5' end of pCAT PrP 3.1 using the -40 primer. This was used to determine the exact amount which was deleted from pCAT PrP 3.5 to produce pCAT PrP 3.1 (also see legend).

**Figure 34.**

A) Schematic representation of pCAT PrP 3.1. This clone contains a 3.1 kb insert. After exonuclease III digestion for 1 min at 37°C only ~ 800 bp of 5' flanking region remained.

B) This was confirmed by double stranded sequencing (with Mn buffer) the clone using the reverse primer (see section 2.4.16 a). The partial sequence is shown. By comparing to the PrP 5' flanking sequence one can see that 402 nucleotides have been removed from pCAT PrP 3.5.

It was not possible, however, to isolate the predicted deletion mutants from other time points. After the 1 min time point, all other time points seemed to undergo some type of recombination or rearrangement event which resulted in most of the insert being lost. This was shown by digesting such clones with BamHI. This produced a linear fragment of approximately 4.5 kb. This meant that the clone only contained an insert of ~ 200 bp. By analysing aliquots of each deletion time point on a 0.8 % agarose gel (figure 33) it could be seen that the exonuclease III deletion reaction had worked, and at the predictable rate of about 450 bp $min^{-1}$ at 37°C. However, when these aliquots were blunt-ended, religated and transformed into E.coli CMK 603 they lost most of their insert, as shown when restriction digests were analysed by agarose gel electrophoresis.

Several variations in the experiment were introduced in an attempt to resolve this problem.

Firstly, the deleted bands were excised directly from the gel, after exonuclease III treatment, and purified before re-ligation and transformation. The rationale behind this was that if any smaller DNA molecules were present after the deletion step then they might preferentially re-ligate, thereby interfering with the expected deletions. By excising the bands directly from the gel this possibility could be avoided. This, however, made no difference in that most clones only had an insert of ~ 200 bp. In addition, when the S1 nuclease step was omitted and the above procedure was repeated the predicted deletions were still not obtained. This problem persisted when new S1 nuclease, exonuclease III, SalI and SphI were used and when a fresh CsCl preparation of pCAT PrP 3.5 was made.

These data suggested that the event which caused the insert size to decrease to ~ 200 bp probably occurred after introduction of exonuclease III treated DNA into E.coli CMK 603.

Transformations of the deletions was then carried out in a different strain (NM554) to determine if a different genotype made a difference. This strain has a *recA* mutation, which eliminates general recombination, but again this was found to make no difference.

Therefore to circumvent this problem, further deletions were made by recloning various restriction fragments of pCAT PrP 3.5 into pCAT basic.

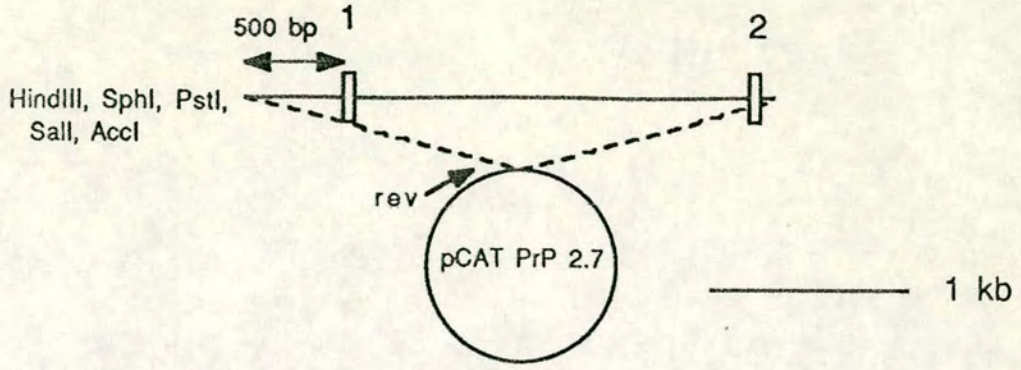
**5.3 Construction of deletion mutants (cont)**

**b) Using Restriction digests**

**(i) pCAT PrP 2.7**

pCAT PrP 2.7 contains a further deletion in the 5' flanking region of the PrP gene. This construct contains approximately 500 bp of sequence 5' to exon 1. pCAT PrP 2.7 was constructed from pCAT PrP 3.1. After exonuclease digestion pCAT PrP 3.1 had the HincII site located at nucleotide number 202 removed (see figures 18 and 20). This meant that a 2.7 kb HincII/SmaI fragment could be easily isolated. This 2.7 kb HincII / SmaI fragment contains ~ 500 bp of 5' flanking region, all of the first intron and 60 bp of exon 2. This fragment was cloned into the blunt ended XbaI site of pCAT basic. Again the orientation was confirmed by double stranded nucleotide sequencing using the reverse primer.

A)




B) AACATGCAATGGCCAATATACTTTCT

Partial nucleotide sequence from the 5' end of pCAT PrP 2.7 using -40 as a primer. This data was used to determine exactly how many nucleotides had been deleted from pCAT PrP 3.1 to produce pCAT PrP 2.7 and also that the insert was in the correct orientation, with respect to transcription, after re-cloning.

**Figure 35**

A) Schematic representation of pCAT PrP 2.7. This clone has a 2.7 kb insert and was made from pCAT PrP 3.1, which had its 5' HincII site destroyed by exonuclease deletion. This allowed a 2.7 kb HincII / SmaI fragment to be cloned into the blunt-ended XbaI site of pCAT basic.

B) Orientation of insert was confirmed by double stranded nucleotide sequencing (with Mn buffer). Partial sequence is shown proving that insert is in the correct orientation with respect to transcription.

5.3 Construction of deletion mutants (cont)

b) Using restriction digests

(ii) pCAT PrP 1.7

pCAT PrP 1.7 has no deletions in the 5' flanking region but it does have a large deletion which removes 1.8 kb of the first intron. Approximately 500 bp of this intron remains. Whether any regulatory elements are found in the introns of the murine PrP gene is not known. However, there is some indirect evidence for the existence of such elements within the intron of the hamster PrP gene. Scott *et al.*, (1989) constructed a minigene which contained hamster PrP promoter attached to the open reading frame of PrP. Transgenic mice which harboured this mini-gene were found not to produce hamster PrP. It might therefore be the case that adequate expression of PrP is dependent on intron sequences. In addition, regulatory elements have been found in the introns of other genes e.g c-fos gene (Lamb *et al.*, 1990). pCAT PrP 1.7 was made by cloning a 1.7 kb SalI / EcoRI fragment from pCAT PrP 3.5 into the SalI site of pCAT basic. After the SalI termini of the fragment had been ligated in, the reaction was blunt-ended to allow the EcoRI termini to be ligated in.

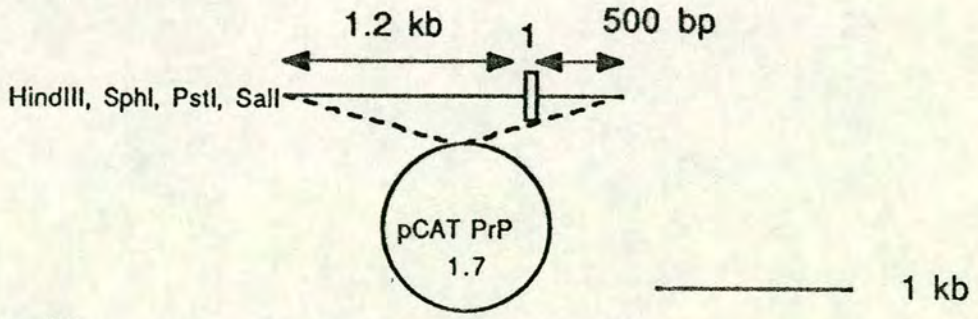



**Figure 36.**

pCAT PrP 1.7 was made by isolating a 1.7 kb SalI / EcoRI fragment from pCAT PrP 3.5. This was cloned into pCAT basic by an initial sticky end ligation of the SalI termini, followed by a blunt end ligation of the EcoRI termini to the SalI termini. This construct contains 1.2 kb of 5' flanking region and ~500 bp of the first intron. No suitable restriction sites were found upstream of the EcoRI site which would have allowed complete removal of the first intron.

## 5.4 Summary of PrP CAT constructs

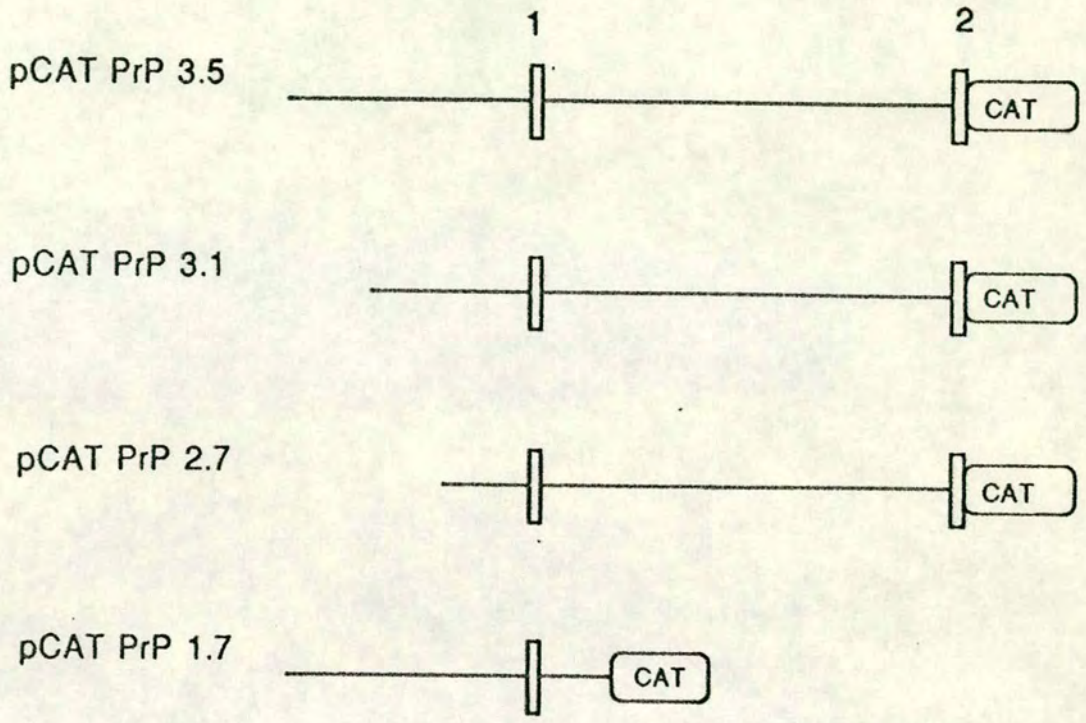



**Figure 37.**

Four constructs were made which contained varying lengths of 5' flanking region and first intron of the PrP gene. All were cloned upstream of the bacterial reporter gene CAT in the correct orientation to promote transcription of CAT. The CAT vector is pCAT PrP basic.

pCAT PrP 3.5 contained a 3.5 kb insert which was composed of 1.2 kb 5' flanking region, exon 1, 1st intron and 60 bp of exon 2.

pCAT PrP 3.1 is a derivative of pCAT PrP 3.5 and was made by exonuclease III deletion. This construct contained ~ 800 bp of 5' flanking sequence, exon 1, 1st intron and 60 bp of exon 2.

pCAT PrP 2.7 is a subclone of pCAT PrP 3.1 and was made by cloning a 2.7 kb HincII/SmaI fragment into pCAT basic. This construct contained ~ 500 bp of 5' flanking sequence, exon1, 1st intron and 60 bp of exon 2.

pCAT PrP 1.7 is a subclone of pCAT PrP 3.5 and was made by cloning a 1.7 kb SalI/EcoRI fragment into pCAT basic. This construct contained 1.2 kb 5' flanking region, exon 1 and ~ 500 bp of the 1st intron.

## 5.5 Calcium phosphate mediated transfection

DNA constructs used for transfections were purified by two rounds of cesium chloride - ethidium bromide gradient centrifugation. This removed any residual RNA, open circular and nicked DNA which could have affected transfection efficiency. The DNA was exhaustively dialysed against tissue culture grade $H_2O$. Transfections using the calcium phosphate method involved mixing the DNA with $CaCl_2$, then vortexing into a phosphate buffer (HEPES buffered saline) to form a fine precipitate. This was then presented to the cells and taken in by endocytosis. 5 μg of the plasmid pSV-β-galactosidase (Promega) was routinely added to the test DNA for every transfection. This plasmid contains the lacZ gene of E.coli which codes for the easily assayable β-galactosidase enzyme. Transcription of the lacZ gene is driven by the SV40 promoter and enhancer. This plasmid was used as a co-transfectant in order to distinguish between differences in transcriptional activity from differences in transfection efficiency. To illustrate this, suppose two different constructs (plasmid A and plasmid B) were being compared for promoter strength and that an internal control was not used. If plasmid A gave an arbitrary promoter strength of 2 and plasmid B gave an arbitrary promoter strength of 4, then this could result in different interpretations. Firstly, it could mean that plasmid B has twice the promoter strength of plasmid A. On the other hand, it could mean that plasmid B had been transfected twice as efficiently. By including an internal control (pSV-β-galactosidase) such anomalies can be circumvented.

Previous studies have shown that PrP is highly expressed in mouse neuroblastoma cell lines such as neuro2a and C1300 (Race *et al.*, 1987, 1988; Butler *et al.*, 1988). It therefore seems sensible to test the PrP CAT constructs in such a cell line. Neuro2a cells were used.

Protocol:

a) The day before transfection the cells were plated out at 1-2 x $10^5$ cells $ml^{-1}$.

b) The media (DMEM + 10 % fcs) was replaced with fresh media 3 hrs before the transfection.

c) 5 μg of test CAT construct was mixed with 5 μg of pSV-β-galactosidase, 37 μl of sterile 2M $CaCl_2$ and the volume was adjusted to 300 μl with sterile $H_2O$.

d) The DNA/$CaCl_2$ was then added, dropwise, to 300 μl sterile HEPES buffer pH 7.1 which was being vortexed during the addition.

e) The precipitate was then allowed to form for 30 mins at room temperature.

f) Just before addition to the cells, the solution was vortexed and then added to the cells in a dropwise fashion. After addition, the plate was gently swirled then replaced at 37°C, 6 % $CO_2$ for 48 hrs.

**5.6 Preparation of cell extracts**

After 48 hrs the cells were removed with a cell scraper and decanted into a 15 ml tube along with the media. The cells were recovered by centrifugation at 1000 x g for 10 mins. The pellet was resuspended in 1ml PBS (Phosphate Buffered Saline) then split into 2 x 500 μl aliquots: one for the CAT assay and the other for the lacZ assay. Both aliquots were spun at 5000 rpm in a microfuge for 5 mins. The pellets were then lysed for 5 mins using 60 μl of a solution of 10mM Tris-HCl pH 8, 2mM $MgCl_2$, 0.5 % Nonidet P40. The samples were then centrifuged in a microfuge for 2.5 mins at 13000 rpm and the supernatant collected. The lacZ

aliquots were usually stored at -70°C until the next day. The CAT assay aliquots were heated at 60°C for 7 mins, to destroy any endogenous deacetylases, then spun for 10 mins at 13000 rpm in a microfuge. The supernatant was used for the CAT assay.

### 5.7 CAT assays

This was done according to the method of Seed and Sheen (1988) and uses butyryl co-A rather than acetyl co-A.
50 µl of cell extract was incubated in a 93 µl reaction volume containing mix A. Mix A was made by mixing the following:

| $H_2O$ | 1M Tris-HCl pH 8 | $^{14}C$ chloramphenicol | 2.5mM butryl co-A |
|---|---|---|---|
| 5 vol | 1vol | 2 vol | 2 vol |

When 43 µl of mix A was added to 50 µl of cell extract the tube was vortexed and incubated at 37°C for 4.5 hrs.
After incubation the reaction mixture was extracted with 200 µl of mixed xylenes by vortexing then centrifugation in a microfuge for 5 mins at 13000 rpm. The butrylated chloramphenicol partitions into the xylene (upper) phase and 180 µl was removed to a fresh tube where it was back extracted with 100 µl of T.E. This back extraction involved vortexing and spinning for 4 mins in a microfuge at top speed. The back extraction was then repeated. This removed any residual un-butrylated chloramphenicol. 150 µl of the top layer was then added to 2 ml of scintillation fluid in a vial and measured in a scintillation counter. Each construct was done in quadruplicate.

## 5.8 β-galactosidase assays

pSV-β-galactosidase was used as an internal control with every transfection in order to standardise transfection efficiencies. The plasmid contains the E.coli lacZ gene (which codes for β-galactosidase) driven by the SV40 promoter and enhancer. β-galactosidase can be measured spectrophotometrically by utilising its ability to hydrolyse the chromogenic substrate ONPG (o-nitrophenyl-β-D-galactopyranoside). This method was an adaption of that of Rosenthal (1987).

For each CAT assay, two β-galactosidase were done and they were set up as follows: 25 µl cell extract, 3 µl 100 x Mg buffer, 50 µl ONPG at 4 mg $ml^{-1}$ and 222 µl 0.1M sodium phosphate buffer at pH 7.5. This was vortexed and incubated at 37°C until a yellow colour developed (1-2 hrs). β- galactosidase assays were also done on a positive control which comprised 0.05 units of β-galactosidase standard (Sigma). Negative controls were mock transfected cells. The reactions were stopped by heating at 75°C for 10 mins and measured in a spectrophotometer at 420 nm. The blank was 1 x Mg buffer, 0.6 mg $ml^{-1}$ ONPG and 0.1M sodium phosphate.

## 5.9 Normalisation of CAT results

To differentiate transfection from transcriptional differences, CAT assays were normalised or standardised against β-galactosidase assays (β-gal). For each DNA construct, at least four separate transfections were performed. In addition to this, at least two different DNA preparations were tested. Normalisation basically involves dividing a CAT assay value (in cpm) with its corresponding β-gal value (in OD 600 nm units). N.B the cell extract from transfected cells was used for both CAT assay and β-gal assay. A mock transfection (no DNA) was also performed to control for endodgenous lysosomal β-galactosidase activity. For every CAT assay performed, two β-gal assays were done and the average value of them determined. This was done as the β-gal assays tended to give more variable results. Normalised results were expressed as a mean (n=4) ± S.E.M.

$$\text{S.E.M} = \frac{SD}{\sqrt{N}}$$

## Tables 5-10.

Tables 5-10 represent data obtained in the normalisation procedures of pCAT PrP 3.5, 3.1, 2.7, 1.7, pCAT control and pCAT basic, respectively. The CAT assays were done in quadruplicate and individual results (expressed in counts per min - cpm) are shown in column a. For each CAT assay two β-galactosidase (beta-gal) assays were done and the average of these calculated. This is referred to as the mean beta-gal value and this is shown in column b. In addition, a β-gal assay was performed for a mock (no DNA) transfection to allow for endodgenous β-gal activity. This control value = 0.79, and when subtracted from the values in column b results in the column c values being produced.

Normalised CAT assay values are produced by dividing CAT assay values (a) by adjusted beta-gal values (c). Then the average of these four normalised CAT assay values (d) is calculated and expressed ± the standard error of the mean (underneath each table).

A summary of these data is shown in figure 38.

## Normalisation of pCAT PrP constructs

Table 5 - Normalisation of pCAT PrP 3.5

| (a) CAT assay (cpm) | (b) Mean beta-gal | (c) Mean beta-gal - control | (d) Normalised CAT assay (a/c) |
|---|---|---|---|
| 160,816 | 1.566 | 0.776 | 207,237 |
| 164,263 | 1.210 | 0.420 | 391,102 |
| 173,902 | 1.450 | 0.660 | 263,487 |
| 175,279 | 1.181 | 0.391 | 448,283 |

Normalised result for pCAT PrP 3.5
327527 ± 55673



Table 6 - Normalisation of pCAT PrP 3.1

| (a) CAT assay (cpm) | (b) Mean beta-gal | (c) Mean beta-gal - control | (d) Normalised CAT assay (a/c) |
|---|---|---|---|
| 193,603 | 1.386 | 0.596 | 324,837 |
| 151,819 | 1.179 | 0.389 | 390,280 |
| 142,248 | 1.348 | 0.558 | 254,924 |
| 152,037 | 1.292 | 0.502 | 302,862 |

Normalised result for pCAT PrP 3.1
318225 ± 28104

Table 7 - Normalisation of pCAT PrP 2.7

| (a) CAT assay (cpm) | (b) Mean beta-gal | (c) Mean beta-gal - control | (d) Normalised CAT assay (a/c) |
|---|---|---|---|
| 81,669 | 1.579 | 0.789 | 103,509 |
| 54,180 | 1.390 | 0.600 | 90,300 |
| 62,829 | 1.335 | 0.545 | 115,284 |
| 104,069 | 1.737 | 0.947 | 109,840 |

Normalised result for pCAT PrP 2.7
104733 ± 5379



Table 8 - Normalisation of pCAT PrP 1.7

| (a) CAT assay (cpm) | (b) Mean beta-gal | (c) Mean beta-gal - control | (d) Normalised CAT assay (a/c) |
|---|---|---|---|
| 45,181 | 1.260 | 0.470 | 96,129 |
| 32,950 | 1.217 | 0.427 | 77,166 |
| 22,180 | 1.092 | 0.302 | 73,443 |
| 44,259 | 1.300 | 0.510 | 86,782 |

Normalised result for pCAT PrP 1.7
83380 ± 5094

## Normalisation of control constructs

Table 9 - Normalisation of pCAT control (positive control)

| (a) CAT assay (cpm) | (b) Mean beta-gal | (c) Mean beta-gal - control | (d) Normalised CAT assay (a/c) |
|---|---|---|---|
| 64,207 | 1.615 | 0.825 | 77,826 |
| 51,458 | 1.524 | 0.734 | 70,106 |
| 71,987 | 1.538 | 0.748 | 76,239 |
| 51,924 | 1.421 | 0.631 | 82,288 |

Normalised result for pCAT control
81614 ± 6713



Table 10 - Normalisation of pCAT basic (negative control)

| (a) CAT assay (cpm) | (b) Mean beta-gal | (c) Mean beta-gal - control | (d) Normalised CAT assay (a/c) |
|---|---|---|---|
| 2,829 | 1.283 | 0.493 | 5,738 |
| 2,587 | 1.136 | 0.346 | 7,476 |

Normalised result for pCAT basic
6607 ± 616

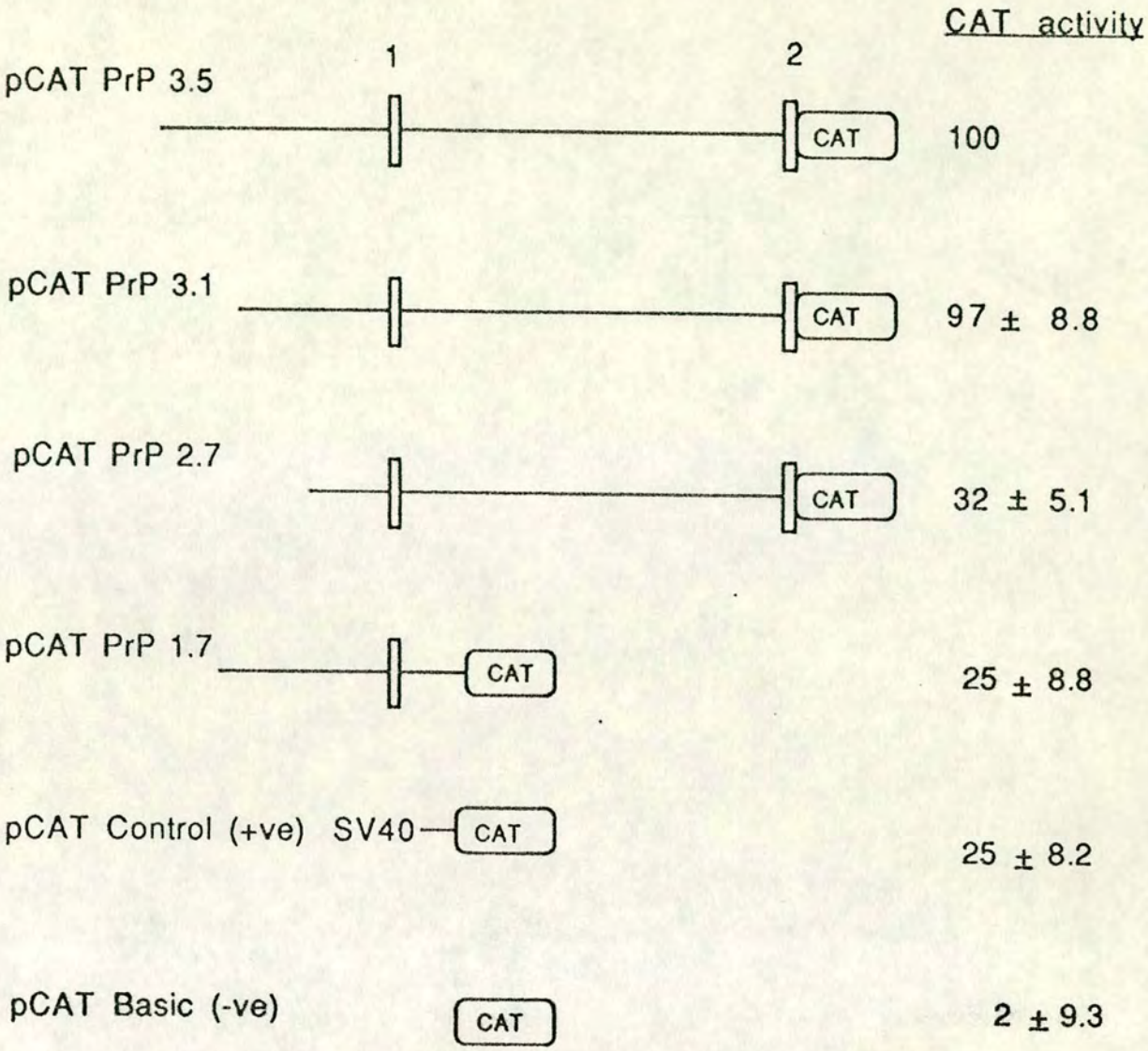



**Figure 38.**

CAT activity of PrP deletion constructs and controls. Constructs containing varying lengths of PrP 5' flanking and intron region were cloned upstream, and in the correct orientation with respect to transcription, of the CAT gene. The positive control encompassed the SV40 promoter and enhancer driving the CAT gene, whereas the negative control consisted of the promoterless plasmid pCAT basic. These constructs were transfected into neuro2a cells by calcium phosphate mediated transfection and after 48 hrs CAT activity was measured. Each construct was done in quadruplicate (except pCAT basic), with CAT activity being normalised to beta-galactosidase activity (see tables 5-10 for raw data). The highest normalised CAT assay value (pCAT PrP 3.5) was arbitrarily designated as having a promoter activity of 100 and all other constructs were expressed as a percentage of this. Each value is an average of four separate transfections ± the standard error of the mean (S.E.M).

## 5.10 Results and Discussion

To identify the regulatory regions responsible for the expression of the mouse PrP gene several 5' deleted fragments of the PrP promoter, and a deletion in the intron, were fused upstream of the CAT reporter gene. The promoter activity of these constructs was then analysed by transfection into murine neuro2a cells and subsequent measurement of CAT activity.

pCAT PrP 3.5 contained 1.2 kb of 5' flanking region, exon 1, all of the first intron and about 60 bp of exon 2 fused upstream of CAT (see figure 32). This construct had the highest promoter strength and, as such, was designated as having an arbitrary promoter strength of 100. All other values were expressed as a percentage of this value. pCAT PrP 3.1 (figure 34a), which contains 800 bp 5' flanking region, exon 1, the entire first intron and about 60 bp of exon 2, had approximately the same promoter strength as pCAT PrP 3.5 (see figure 38).

pCAT PrP 2.7 is the same as pCAT PrP 3.1 except that it has only 500 bp of sequence 5' to the first exon. This deletion results in a greater than 3-fold reduction in CAT activity when compared to pCAT PrP 3.5 (figure 38). This suggests that the PrP promoter possesses a positive regulatory element which resides within the deleted region (~ -500 to -800). pCAT PrP 1.7 has no deletions in the 5' flanking region but it does have a large deletion in the first intron. This construct contains 1.2 kb 5' flanking region and approximately 500 bp of intron 1. pCAT PrP 1.7 shows, approximately, a 4-fold reduction in CAT activity when compared to pCAT PrP 3.5 (see figure 38). This analysis has shown that there may be a positive regulatory element in the first intron of the PrP gene. The level of CAT activity from pCAT PrP 1.7 is comparable with that of the positive control (pCAT control) and is approximately 12.5-fold higher than the negative control (pCAT basic). In order to determine the net effect of the first intron on the expression of CAT, one would have to delete it entirely. As it is, pCAT PrP 1.7 contains ~ 500 bp of intron.

Probably the simplest way of deleting all of the first intron would be to amplify the promoter region by PCR and then reclone it into pCAT basic. Such a construct would then be re-tested in a CAT assay after transfection.

A more detailed analysis of the PrP promoter region could be gained by creating more deletions in the 5' flanking region. This could be done by using Sph I/ SalI digested pCAT PrP 2.7 as substrate for exonuclease III. This would produce unidirectional deletion within the 500 bp region which is 5' to exon 1. Most of the potential transcription factor binding sites are located in this 500 bp region. The effect of their deletion could therefore be ascertained.

The positive control used in these transfection experiments was pCAT control. This construct contained the CAT gene which was driven by the SV40 promoter and enhancer sequences. This construct gave approximately 4-fold less CAT expression when compared to pCAT PrP 3.5 but approximately the same levels of expression as did both pCAT PrP 3.1 and pCAT PrP 2.7. Other workers have also found that the SV40 promoter is not particularly strong in neuronal cell lines. For example, Donis *et al.*, (1993) compared the strength of several different promoters when linked to CAT and transfected into PC12 cells. Like neuro2a cells PC12 cells are a neuronal cell line but are derived from rat rather than from mouse. In that study the SV40 promoter was shown to be weaker than the long terminal repeat of the Moloney murine leukaemia virus (MMLV), the Rous sarcoma virus (RSV) promoter, the beta-actin promoter and the simian cytomegalovirus (SCMV) promoter. Indeed, the SV40 promoter directed only five times the amount of CAT activity when compared to cells which were not transfected. Similar results were found with an immortalised neuronal cell line called HT4.

For a negative control in this study, the promoterless pCAT basic plasmid was used. This control gives one an estimation of the `background noise' of the basic plasmid. The relatively small CAT activity from pCAT basic (8% of pCAT control) is probably due to spurious promoter elements located in the vector.

This assumption is supported by experimental evidence from the no DNA control (mock transfection) which gave an average cpm of 226 (n = 2) in the CAT assay (data not shown). This means that there is very little endodgenous CAT activity in the neuro2a cells.

An important assumption which is often made in studies involving reporter gene constructs is that transcriptional activity is directly proportional to the activity of the reporter enzyme. This assumption is generally true (Alam and Cook, 1990), although it need not necessarily hold in every case. For example, pCAT PrP 1.7 has 1.2 kb of 5' flanking region, exon 1 and about 500 bp of intron. This intron is only flanked by one splice site at the exon 1 / intron border and hence this could affect RNA stability or its ability to be processed correctly. This may affect the levels and possibly the activity of the CAT enzyme. Therefore a more direct approach which could have been used instead, or additionally as a control, would be to measure CAT mRNA directly rather than CAT enzyme activity. This could have been done by quantitative RT-PCR, primer extension or by ribonuclease protection.

## 6.1 Introduction

Primer extension is a technique which can be used to determine the 5' termini of mRNA. This by definition determines where transcription starts and can be mapped precisely on the DNA if nucleotide sequence of this region is available. The principle of the technique is that a radioactively labelled primer, either a single stranded restriction fragment or an oligonucleotide, corresponding to the gene of interest is hybridised to RNA. Reverse transcriptase is then used to extend this annealed primer to produce a cDNA. This extended product is then analysed on a denaturing polyacrylamide gel with e.g M13 sequence as size markers. The extension reaction should terminate at the extreme 5' end of the RNA.

The primer extension reaction is more sensitive and gives rise to less artefacts with poly ($A^+$) RNA rather than total RNA. Further artefacts can be prevented by designing primers which are as close to the 5' terminus of the RNA as possible (preferably within 100 nucleotides of the 5' terminus).

Primer extension analyses have previously been performed on the PrP gene using human (Puckett *et al.*, 1991), hamster (Basler *et al.*, 1986) and mouse (Westaway *et al.*, 1987) RNA.

Puckett *et al.*, (1991) analysed the human PrP gene by primer extension using a 30 mer as a primer and poly ($A^+$) RNA and found that transcription initiated at one site. This was located approximately 136 bp upstream of the 5' splice site. In contrast, when Basler *et al.*, analysed the hamster PrP gene by primer extension, using a 35 mer as a primer and poly($A^+$) RNA, multiple transcriptional start sites were identified. These were found between 55 and 80 nucleotides upstream of the 5' splice site. In 1987 the promoter region of the murine PrP gene was uncloned.

However, Westaway *et al.*, (1987) found that murine PrP RNA had heterogenous 5' termini when analysed by primer extension. In this case a 30 mer primer and total RNA were used. Due to the promoter region being uncloned, however, the multiple 5' termini of the RNA could not be mapped to the promoter region of the gene.

The group of experiments described below reproduce and expand those of Westaway *et al.*, (1987), but also contain several important modifications. These include the use of a different oligonucleotide as a primer and of poly ($A^+$) RNA instead of total RNA, which often gives unacceptably high levels of premature terminations. Also the transcriptional start sites could be mapped directly to the promoter region as this has been cloned and its nucleotide sequence determined (see chapter 4).

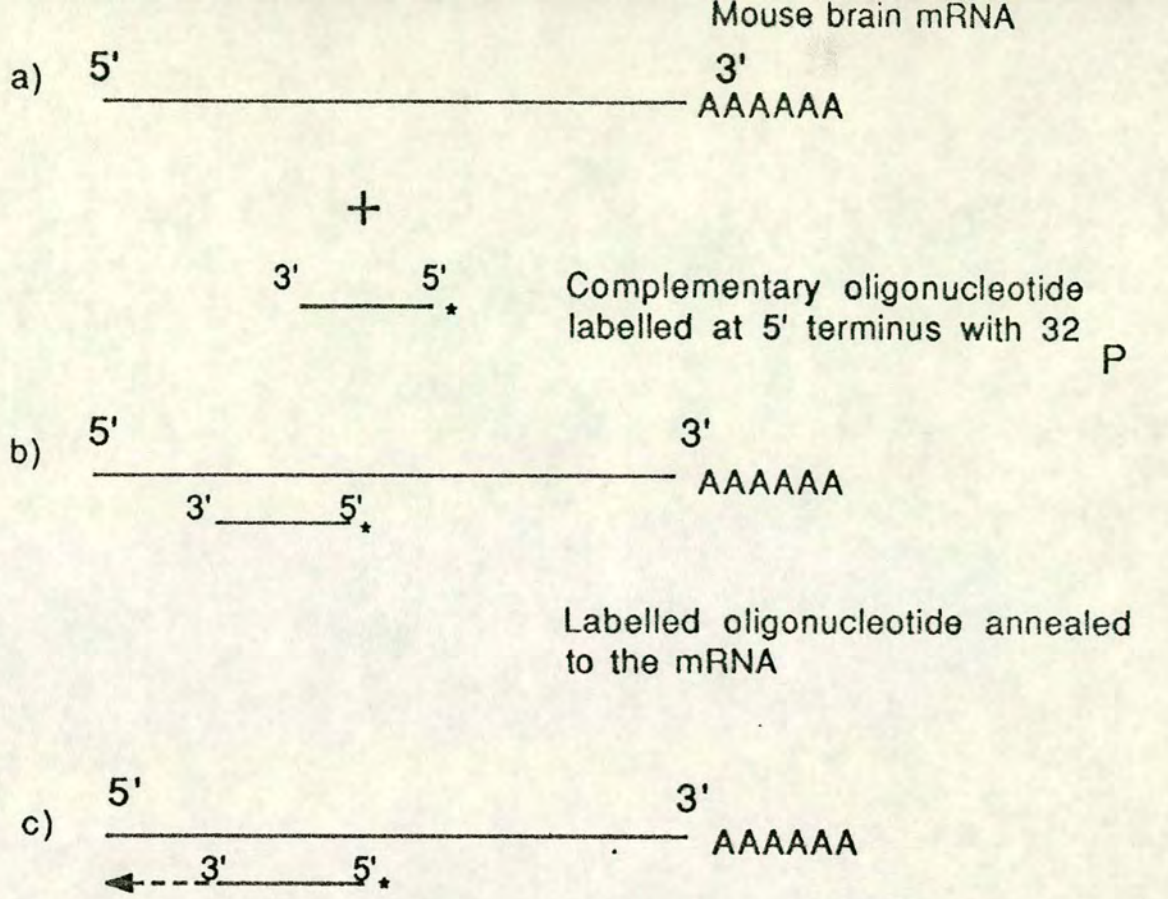



**Figure 39.**

Schematic representation of primer extension.

Approximately 20,000 cpm of exon 2 specific oligonucleotide (designated G3206) was annealed to five micrograms of brain mRNA derived from a VM/Dk mouse (henceforth referred to as VM). This was performed by heating to 90°C for 10 mins and then allowing to cool to room temperature. The primer was then extended to the 5' terminus of the mRNA by incubating at 37°C with Moloney murine reverse transcriptase and dNTPs for 1 3/4 hrs. The extension products were analysed by denaturing polacrylamide gel electrophoresis. M13 sequence was used as size markers. Asterix denotes radioactive label and broken arrows represent the extension reaction.

## 6.2 Isolation of total RNA

Total RNA was initially purified from a frozen brain (0.457 g) of a VM (Sinc p7) mouse and then later scaled up to produce larger amounts of RNA. The brain was ground in a small volume of liquid nitrogen using a mortar and pestle. The powder was added to 9 ml of RNAzol (Biogenesis Ltd) and then mixed by vigorous vortexing. 900 µl of chloroform was added to the homogenate and again vigorously vortexed. The homogenate was then incubated on ice for at least 5 mins and then spun at 12000 x g for 15 mins at 4°C. After centrifugation the suspension separated into two phases: the lower blue phase and the upper colourless phase. The upper phase was transferred to a fresh tube. An equal volume of isopropanol was added, mixed, and stored at 4°C for 15 mins. After this the tube was spun at 12000 x g for 15 mins at 4°C to recover the RNA. The pellet was then washed in 70 % ethanol and spun at 7500 x g for 8 mins. The pellet was resuspended in 500 µl of DEPC treated $H_2O$ and stored at -20°C. Approximately 470 µg of total RNA was recovered from 1 brain.

## 6.3 Isolation of mRNA

There are several methods which can be used for the isolation of mRNA, all of which utilise the poly($A^+$) region which is found at the 3' region of most eukaryotic mRNAs.

The system described here relies on magnetically separating mRNA from other RNA species. The polyAtract mRNA isolation system (Promega) was used. This system involves annealing a biotinylated oligo(dT) primer to the 3' poly($A^+$) region of the mRNA. Streptavidin paramagnetic particles are then added to the annealed mix and these bind to the biotin.

On application of a magnetic force to the side of the tube, the strepavidin paramagnetic particles which are coupled to the oligo(dT)/biotin mRNA complex are pulled to the side of the tube. These are washed several times and the mRNA is then eluted.

Protocol

a) 1 mg of total RNA in a volume of 500 µl $H_2O$ was heated at 65°C for 10 mins. 3 µl of the biotinylated-oligo(dT) probe and 13 µl of 20 X SSC were then mixed into the RNA and allowed to cool to room temperature. Meanwhile, the streptavidin paramagnetic particles were resuspended by gently flicking the bottom of the tube and the captured onto the side of the tube using a magnet. The particles were washed three times in 0.5 X SSC (0.3 ml per wash) and collected each time using the magnet. The particles were resuspended in 0.1 ml of 0.5 X SSC.

b) Then the entire annealing reaction was added to the washed streptavidin paramagnetic particles and incubated for 10 min. The particles were then recaptured using the magnet and washed four times in 0.1 X SSC (0.3 ml per wash).

c) To elute the mRNA, 100 µl of $H_2O$ was used to resuspend the captured particles. After recapturing the particles the eluate was transferred to a fresh tube. The elution procedure was repeated using 150 µl of $H_2O$. The two eluates were then pooled.

Typically, 12 µg of poly ($A^+$) was isolated from 470 µg of total RNA, which was the equivalent of one mouse brain. This is consistent with approximately 1-2 % of total RNA being composed of mRNA (Sambrook *et al.*, 1989).

### 6.4 Northern blot of PrP mRNA

A Northern blot was carried out to check the integrity of the PrP mRNA.

#### a) Electrophoresis of RNA

A 100 ml gel was made with 72 ml $H_2O$, 18 ml formaldehyde, 1 g agarose and 10 ml of 10 x RNA running buffer. An equal volume of formamide sample buffer was added to 1 μg of brain mRNA and heated to 75°C for 5 mins and then quickly cooled on ice. 0.25 vol of Ficoll-dye-EDTA was added, mixed and then loaded onto the gel immediately. The gel was run at 15 volts overnight.

#### b) Northern transfer

After electrophoresis the gel was washed in 500 ml of $H_2O$ for 10 mins to remove traces of the formaldehyde from the surface of the gel. The set up for the Northern blot was essentially the same as for a Southern blot, except that the transfer buffer for the Northern blot was 20 x SSC. The RNA was blotted onto Hybond-N overnight. After transfer the membrane was washed in 2 x SSC, dried and fixed by baking at 80°C for 2 hrs.

#### c) Prehybridisation of Northern blot

The membrane was prehybridised in 50 ml of Hybond-N prehybridisation buffer at 65°C overnight in a shaking water bath.

d) Hybridisation and washing of Northern blot

A 2 kb PstI fragment (see figure 40), encompassing the ORF of mouse PrP gene, was labelled with $^{32}P$ by random priming (see section 2.4.11 a) and used to probe the Northern blot. The 2 kb PstI fragment was derived from exon 3 of the mouse PrP gene - see diagram below. The probe was boiled for 7 mins and then added to the hybridisation solution and incubated overnight at 65°C.

The membrane was washed twice for 15 mins in 0.1 x SSC, 1 % SDS at 65°C. A 2.4 kb band was evident from the resulting autoradiograph and little RNA degradation was detected (see figure 41). The primer extension experiments were then performed.

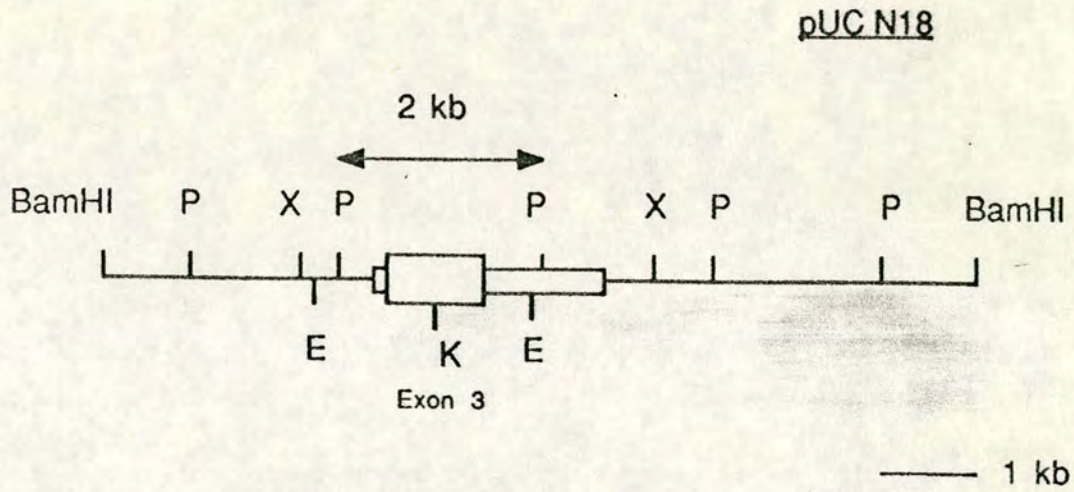



**Figure 40.**

pUC N18 is a 8.8 kb genomic clone of a NZW mouse PrP gene (provided by D. Westaway). This clone encompasses exon 3 and surrounding region. The entire open reading frame is indicated with a large box. The narrower open box designates the untranslated regions of exon 3. The single straight line at the 5' end of the insert represents part of intron 2, whereas the same line at the 3' end represents 3' flanking sequences. The 2 kb PstI fragment used as a probe in the northern blot is indicated. Restriction sites: P, PstI; E, EcoRI; K, KpnI and X, XbaI.

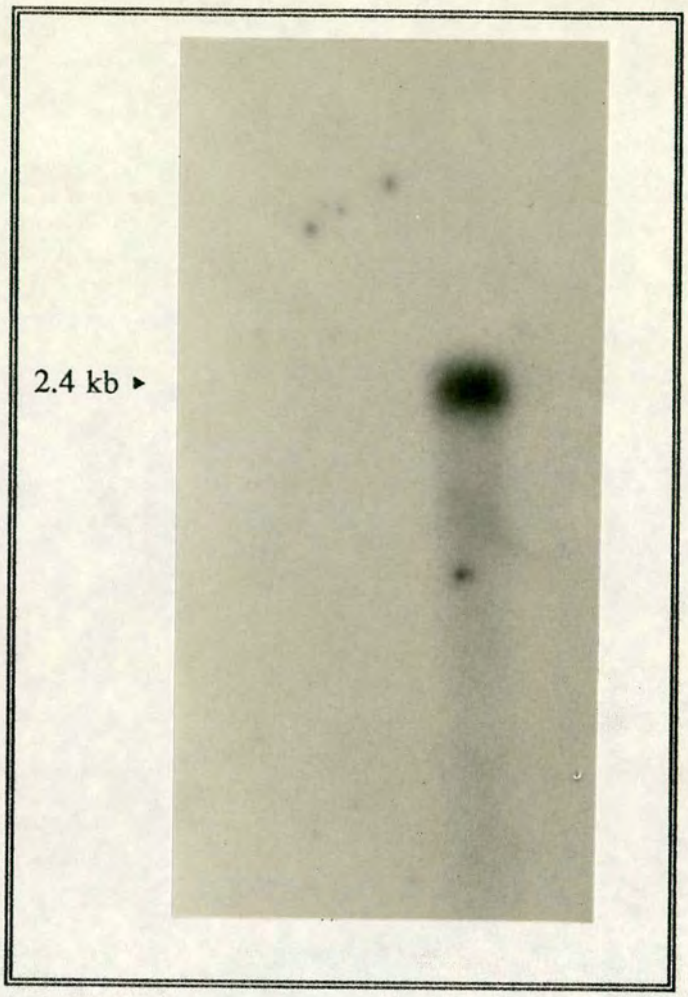



**Figure 41**

Northern blot of PrP mRNA from a VM mouse.

One microgram of poly ($A^+$) brain RNA from a VM mouse was electrophoresed overnight on a 1 % agarose gel containing formaldehyde. After electrophoresis the gel was washed in $H_2O$ for 10 mins and blotted overnight onto Hybond N, using 20 x SSC as the transfer buffer. The membrane was then baked at 80°C for 2 hrs to fix the RNA to the membrane. A 2 kb PstI restriction fragment encompassing exon 3 (see figure 40) was radioactively labelled with $^{32}P$, using random priming, and hybridised to the RNA at 65°C overnight in 0.5M sodium phosphate, 7 % SDS. The membrane was then washed twice, for 10 mins, in 0.1 x SSC, 1 % SDS at 65°C and autoradiographed overnight at -70°C. A single band of approximately 2.4 kb was evident. Gibco BRL RNA markers were used as size standards (not shown) and ranged from 0.24 to 9.5 kb.

## 6.5 Primer extension

### a) Design of oligonucleotide primer and its location on mouse cDNA sequence

By referring to the published mouse cDNA sequence (Locht *et al.*, 1986) an oligonucleotide was designed which corresponded to the 5' terminus of the cDNA sequence. The position of this oligonucleotide (designated G3206) with respect to the cDNA sequence is shown below. Oligonucleotide G3206 was designed to bind as close to the 5' terminus as possible, based on the information which was available at the time. This reduces the possibility of the reverse transcriptase either becoming dissociated from the RNA or stopping at a region of high secondary structure (Sambrook *et al.*, 1989). The first 9 nucleotides of the cDNA sequence (Locht *et al.*, 1986) were not found in the exon 2 sequence which is presented in chapter 4 (figure 15) and hence were ignored in the design of G3206. It seemed likely that this 9 bp region was a cloning artefact, and indeed it was not found in the recent sequence analysis of exon 2 by Westaway *et al.*, (1994 c).

G3206

5' 1 AATTCCTTCA GAACTGAACC ATTTCAACCG AGCTGAAGCA

41 TTCTGCCTTC CTAGTGGTAC CAGTCCAATT TAGGAGAGCC

81 AAGCAGACTA TCAAGTCATG GCGAACCTTG 3'

**Figure 42.**

Location of G3206 on PrP cDNA.

The first 110 nucleotides (from 2.1 kb) of the mouse PrP cDNA sequence are shown. The ATG highlighted indicates the translational start codon found in exon 3. Oligonucleotide G3206 is derived from exon two sequence and corresponds to position 10 - 30 on the cDNA sequence. The target sequence for the primer is underlined.

G3206 5'-CGGTTGAAATGGTTCAGTTCT-3'

**6.5 b) Labelling of oligonucleotides**

100 ng of oligonucleotide was labelled with $^{32}P$ dATP and polynucleotide kinase as described in Materials and Methods. After labelling, the DNA was phenol:chloroform extracted. Unincorporated label was removed by 2 ethanol precipitations in the presence of 2M ammonium acetate.

**c) Annealing of oligonucleotide to mRNA**

Approximately 1 - 2 x $10^4$ cpm (1pmol) of labelled oligonucleotide primer was mixed with 5 µg of poly ($A^+$) RNA in the presence of 50mM Tris HCl pH 8.3, 75mM KCl, 10mM DTT, 3mM $MgCl_2$ in a total volume of 10 µl. The oligonucleotide was allowed to anneal to the RNA by heating at 90°C for 10 mins then leaving to cool to room temperature.

**d) Extension reaction**

The annealed primer was extended by adding 0.5 µl of each dNTP and 1 µl of Moloney murine leukaemia virus reverse transcriptase and incubating at 37°C for 1 hr. Another 1µl of enzyme was added and the reaction was continued for another 45 mins. The cDNA product was then ethanol precipitated, washed in 70 % ethanol and then resuspended in 4 µl of loading buffer.

**e) Polyacrylamide gel electrophoresis**

Samples were heated to 90°C for 10 mins before being loaded onto a 6 % sequencing gel. M13 sequencing reactions were used as size markers. The gel was run at about 1500 volts until the first dye migrated into the buffer. The gel was then treated in the same way as a sequencing gel (see section 2.4.16 c) except that it was exposed to x-ray film overnight at -70°C, instead of overnight at room temperature.

## 6.6 Results

### a) Primer extension of mouse brain mRNA with oligonucleotide G2306

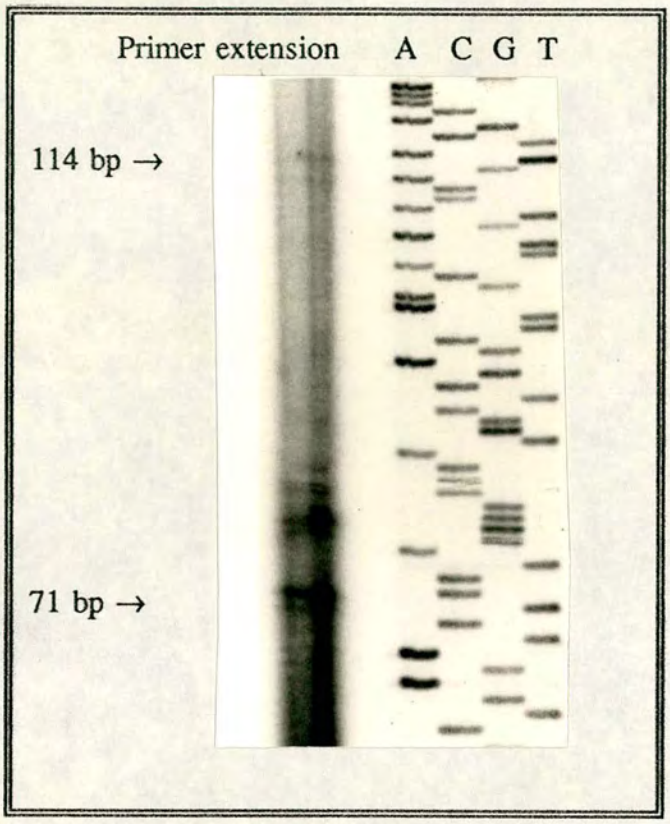



**Figure 43.**

Primer extension analysis of brain mRNA.

There are at least 10 extension products visible on the gel (left hand side of photo). M13 sequence is shown on the right hand side of photo and is used as size markers. There total sizes of the extension products are 70, 71, 76, 77, 78, 80, 81, 83, 111 and 114 bp in length. Some of the extension products are marked with arrows and their sizes shown. By subtracting 21 bp to account for the primer then the actual extensions from the 3' end of the primer are 49, 50, 55, 56, 57, 59, 60, 62, 90 and 93.

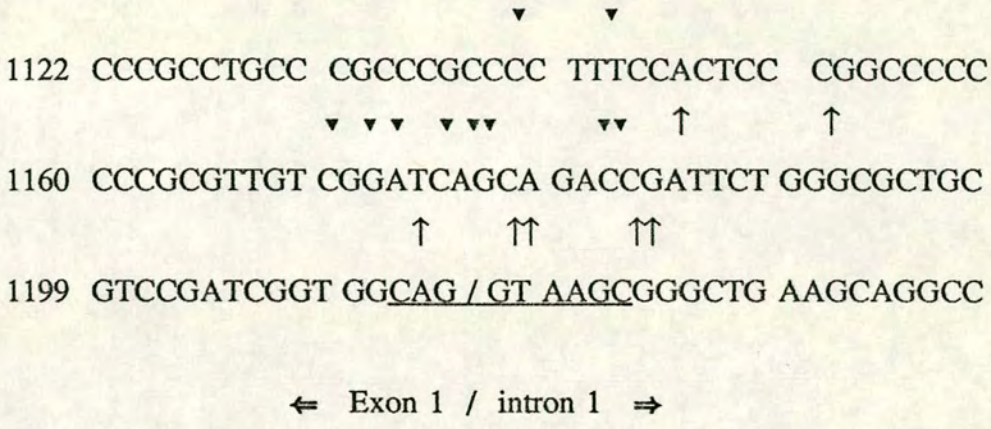



**Figure 44.**

Position of transcriptional start sites in mouse PrP promoter.

Transcriptional start sites (▾ ) were deduced from a primer extension assay using an exon 2 specific oligonucleotide as a primer. Based on the sequence data of exon 2, presented in chapter 4 (figure 15), this primer was predicted to bind to a target sequence which is 17 bp from the intron 1/ exon 2 splice junction. By taking this into account, and the 21 bp which constituted the oligonucleotide, the remainder of the total extension length must correspond to sequence 5' to the exon 1/ intron 1 splice site (represented with a slash and underlined). The complete sequence of the promoter is given in chapter 4 (figure 20 ).

↑ represent start sites for transcription which were recently deduced by Westaway *et al.*, (1994 c).

## 6.8 Discussion

When a PrP exon 2 specific oligonucleotide is used in the primer extension assay multiple extension products are observed. At least ten different extension products were identified and these may correspond to transcriptional start sites within a 44 bp region of PrP 5' flanking sequence. These extension products have variable band intensities and this may reflect the frequency with which a particular site is utilised. From figure 43, one can see that the extension product of 71 bp (second band from bottom) is the most intense, closely followed by the product of 77 bp. Therefore these two sites, corresponding to nucleotide 1134 and 1140 of PrP 5' flanking region, probably represent major start sites for transcription in the brain. The remainder of the extension products are less intense and therefore will represent start sites which are less frequently used for initiating transcription.

Initiation of transcription in the PrP gene of hamster (Basler *et al.*, 1986) and sheep (Westaway *et al.*, 1994 b) also occurs at multiple sites. This, however, is not the case in human PrP gene where transcription starts at one site (Puckett *et al.*, 1991). The significance of this difference, if any, is not known at present. However, it appears that transcriptional initiation at multiple sites, or at one site, bears no correlation as to whether disease in that animal can actually occur. The effects of multiple transcriptional start sites may be more subtle, however, in that while they do not cause disease *per se*, they may still influence how PrP gene is regulated. Assuming that the sequence CAG/GTAAAG (see figure 44) is the correct splice site consensus signal, then the length of exon 1 may vary from 33 bp to 75 bp, depending on which transcriptional start site is utilised. The inclusion, or exclusion, of this stretch of DNA into the PrP transcript may be of importance if this sequence performs a regulatory function. Work with other genes, such as thymidine kinase, has shown that 5' non-coding sequences can have a negative effect on translational efficiency, of the corresponding mRNA, if such sequences have the capacity to form secondary structures (Pelletier and Sonenberg, 1985).

Genes such as c-myc (Piecheazyk *et al.*, 1985), human histone H3 (Morris *et al.*, 1986) and ferritin (Goosen *et al.*, 1990) also have regulatory elements residing in their 5' untranslated regions (UTRs). Whether PrP gene has regulatory elements associated with its 5' UTR is not known, however it is possible that differential use of transcriptional start sites could affect the regulation of such elements.

To determine if the putative splice sites (see above and figure 20) are correct one could have initially performed RT-PCR on PrP mRNA to produce a full-length cDNA. Then, one could sequence the 5' region of this cDNA and compare it to the sequence of exon 1(see figure 44) and exon 2 (see figure 15), allowing for where splicing was expected to occur.

The discrepancies (figure 45) between the mapping of transcriptional start sites, onto the promoter, in this study when compared to the study of Westaway *et al.*, (1994 c) may be attributable to several factors. Firstly, in chapter 4 it was made clear that there were differences in the nucleotide sequence of PrP 5' flanking sequence and also a one nucleotide difference in the sequence of exon 2. This therefore would have introduced inaccuracies in the designation of where transcription starts on the PrP promoter. Another factor which may have produced differences between the two studies is that different mouse strains were used for the isolation of the RNA. In this study, VM mice were used, whereas Westaway *et al.*, (1987) used I/Ln mice. Despite these differences, there is good agreement between both studies and they confirm multiple transcriptional start occur within the mouse PrP gene and that most of them are located within, approximately, a 44 bp region (see figure 44).

Due to the high secondary structure which would be expected within the G C rich PrP promoter, one must interpret the data presented in this chapter with caution. As well as mapping true 5' termini of mRNA the primer extension technique also maps sites where the reverse transcriptase pauses or stops, often due to high secondary structure. Therefore some of the extension products seen in figure 43 may be artefacts. This is

also a criticism of the work of Westaway et al., (1987) as similar potential problems would have existed.

With hindsight, it may have been better to use a more thermostable enzyme rather than Moloney murine leukaemia virus reverse transcriptase which has an optimum polymerisation temperature of 37°C. An enzyme such as TET-z polymerase (Amersham) has reverse transcriptase activity and can function at 70°C . At this temperature much of the predicted secondary structure associated with the PrP promoter would be eliminated and hence a more reliable primer extension result would be obtained. In addition, the primer extension results could have been confirmed with S1 nuclease mapping.

## Chapter 7 Summary and Discussion

To understand if PrP gene regulation is altered in a diseased animal one must first address how the PrP gene is regulated in healthy animals. At present there is virtually nothing known about the DNA sequences and trans-acting factors which control the expression of PrP in different tissues and during development. It is also not known if the regulation of PrP gene expression is affected during disease.
This study has made an initial investigation into PrP gene regulation, in uninfected mice, by analysing the promoter region and the first intron using several different techniques. Each individual results chapter has its own discussion section attached to it. This chapter aims to recapitulate the data obtained in this project and to suggest ideas for further experiments based on this data.

### 7.1 Summary

Firstly, a subclone was isolated from a mouse PrP cosmid clone, cos 6.I/LnJ-4, (Westaway *et al.*, 1991). By nucleotide sequence analysis this subclone, pT18 PrP 3.5, was shown to contain part of exon 2 of the PrP gene. This was based on its high homology to the published PrP cDNA sequence (Locht *et al.*, 1986). Sequence analysis from the opposite (5') end of the subclone revealed that approximately 1.2 kb of PrP 5' flanking sequence was present as was the first exon. This region showed high homology to the hamster PrP promoter but lower homology to the human and sheep PrP promoters. The presence of splice site consensus sequences bordering exon 1 and exon 2, confirm earlier work which suggested that the mouse PrP gene has two non-coding exons separated from each other by an intron of undisclosed size (Westaway *et al.*, 1991).
Secondly, studies were conducted on the promoter activity of various subclones of the PrP gene. A fragment containing 1.2 kb of the 5' flanking region, exon 1, intron and part of exon 2 was cloned upstream of the reporter gene CAT, in the reporter vector pCAT basic.

When transfected into neuro2a cells this construct (pCAT PrP 3.5) was found to direct high expression of CAT. This level of expression was greater than 3-fold higher than the positive control. The positive control contained the SV40 promoter and enhancer linked to the CAT gene. A deletion was made in pCAT PrP 3.5 which removed 400 bp at the 5' end so that 800 bp of 5' flanking region remained. This construct was called pCAT PrP 3.1 and its promoter strength was approximately the same as pCAT PrP 3.5. A further deletion in the 5' flanking region of the PrP gene removed 300 bp and this left 500 bp of promoter region. This construct, pCAT PrP 2.7, showed a marked decrease (3-fold) in promoter activity when compared to pCAT PrP 3.5. This suggests that a positive regulatory element resides in the PrP promoter at position -500 to -800. To determine if the first intron plays a regulatory role, a large deletion was introduced into this region. pCAT PrP 1.7 contained the full 5' flanking region (1.2 kb) but only ~500 bp of the first intron remained. This construct showed a 4-fold reduction in CAT activity when compared to pCAT PrP 3.5.

Thirdly, the transcriptional start sites were determined from the PrP gene using the primer extension technique. An oligonucleotide was synthesised, based on the published mouse cDNA sequence (Locht *et al.*, 1986). This oligonucleotide was labelled at the 5' termini and annealed to mRNA derived from the brain of a VM (*Sinc* p7) mouse. The annealed primer was then extended using MMLV reverse transcriptase and dNTPs. After this extension the reactions were ran on a 6% polyacrylamide sequencing gel and the length of the extended products calculated. M13 sequence was used as size markers. Using this technique multiple transcriptional start sites were identified within a 44 bp region of the PrP promoter.

## 7.2 Function of multiple transcription start sites and the possibility of RNA splicing.

That transcription starts at multiple sites in the mouse PrP gene is probably attributable to its lack of a TATA box in the promoter region. The TATA box controls accurate initiation of transcription (see Smale and Baltimore, 1989, for review) and in genes which lack it, often housekeeping genes, multiple sites for transcriptional initiation are common (Mitchell *et al.*, 1986; Herrick *et al.*, 1993). As mentioned earlier, there is at least one instance whereby differential utilisation of transcription start sites provides a mechanism whereby gene expression can be regulated. This occurs in the α-amylase gene; one start site is utilised when the gene is expressed in the pancreas and another is utilised when it is expressed in the salivary gland (Young *et al.*, 1981).

Whether the use of differential transcriptional start sites plays a role in the regulation of the PrP gene, perhaps under certain circumstances or in certain cell types, is not known. Depending on the transcriptional start site, exon 1 could be between 32 - 75 bp in size. If regulatory sequences exist in exon 1, then their inclusion, or exclusion, could be dependent on which start point for transcription is utilised.

There are examples of regulatory elements residing in the 5' untranslated regions of other genes. For example, an iron responsive element is located in the 5' leader sequence of the ferritin gene (Goosen *et al.*, 1990). In some genes the rate of mRNA decay has been shown to be controlled by their 5' leader sequence such as those which are found in the human histone H3 mRNA (Morris *et al.*, 1986) and in the c-myc mRNA (Piechaczyk *et al.*, 1985; Rabbits *et al.*, 1986). The human c-myc gene consists of three exons. Exons 2 and 3 code for the protein, whereas exon 1 is noncoding. Two forms of the gene can exist; normal and truncated (i.e lacking exon 1). Truncation can occur by a translocation event in a manner analogous to that in Burkitts lymphoma. Normal c-myc mRNA has been shown to have a half life of approximately 30 mins. In contrast, the truncated c-myc mRNA has a substantially

increased half life, suggesting that exon 1 of c-myc is involved in mRNA stability (Rabbits *et al.*, 1985).

It is therefore conceivable that the upstream exons of PrP may also fulfil such a regulatory role and that this regulation may be adversely affected during disease. For example, suppose that there are two mRNA species; one disease specific (message 1) and the other which is found in uninfected animals (message 2). Now, previous studies have shown (Oesch *et al.*, 1985; Chesebro *et al.*, 1985) that there are no detectable differences in PrP mRNA levels between infected and uninfected animals. This has only been determined, however, at one time point and as yet no-one has attempted to determine PrP mRNA differences between infected and uninfected animals over a time period. This means that if such stability differences do exist then they might not have been detected. Such mRNA stability differences could be mediated by either differential transcription start site usage or by alternatively spliced PrP forms. The repercussions of this are that if message 1 is more stable, it would have more opportunity to be translated in comparison to message 2. This hypothesis becomes more relevant when one considers the published data which shows that the levels of normal PrP protein inversely correlate with the length of incubation period (Prusiner *et al.*, 1990). Indeed, recently it has been shown that overexpression of PrP, involving no scrapie challenge, can result in a neurodegenerative disease in mice (Westaway *et al.*, 1994a).

Mouse (Westaway *et al.*, 1987, 1991) and sheep PrP genes (Westaway *et al.*, 1994 b) have two small exons which are located far upstream from the open reading frame exon. In contrast, hamster (Basler *et al.*, 1987) and human (Puckett *et al.*, 1991) PrP genes only have one 5' exon. The relevance of these differences on PrP gene expression is not known. In mouse and sheep, therefore, there exists the possibility that PrP gene expression may be under an additional level of control: RNA splicing. Gene regulation by means of RNA splicing of noncoding exons, in other genes, is well documented. Brain derived neurotrophic factor (BDNF) gene, for example, has four upstream noncoding exons which are separated from the open reading frame

exon by a large intron (> 16.5 kb). By alternative splicing events several different messages are transcribed which are expressed predominantly in certain tissues. Messenger RNAs which contain exon I, II and III have been found to be predominantly expressed in the brain, whereas exon IV containing transcripts were predominantly expressed in the heart and lung (Timmusk *et al.*, 1993). In the case of murine PrP, however, it is not known if splicing plays a similar role. In a recent study, though, Westaway *et al* (1994 c) showed, by PCR analysis, that in uninfected neonatal and adult brain exon 2 appears to be present in most mRNAs. As mentioned earlier, however, this analysis was not carried out in different tissues or in scrapie infected animals. Therefore in some circumstances RNA splicing may have a role in PrP gene expression. Perhaps a better way of addressing whether splicing of PrP occurs, would be to conduct *in situ* hybridisation analysis.

*In situ* hybridisation is a method which can be used to detect mRNA molecules *in situ* i.e in tissue sections. Firstly, one would determine if splicing of exon 2 occurs. Using sequence data, obtained in this study, for exon 1 and exon 2, specific oligonucleotides could be synthesised for each exon. As a control an oligonucleotide from exon 3 could also be used. These oligonucleotides would be radioactively labelled and used to probe tissue sections from various organs e.g brain, heart, kidney, spleen and liver. If RNA splicing does not occur one would expect that each probe would hybridise to the same tissue and to the same cell types within that tissue. If RNA splicing does occur then one would expect the exon 1 and 2 probes to hybridise to different cellular locations or to different cell types. Of course the same protein would be translated whether splicing occurred or not as the entire open reading frame is found in exon 3. However, the different messages transcribed may be regulated differently.

One assumption that would be made with such experiments is that exon 1 is always included in the transcribed message. From the primer extension results presented in chapter 6, this seems to be the case, at least, with PrP mRNA isolated from brain. It is worth bearing in mind however that there are examples in other genes were multiple promoters are found within one gene, leading to the production of different mRNA isoforms. Examples of such genes include alcohol dehydrogenase (Benyajati *et al.*, 1983), a fos responsive gene called Fit-1 (Bergers *et al.*, 1994) and the myosin light chain gene 1/3 (Garfinkel and Davidson, 1987). The different mRNA isoforms transcribed from genes with multiple promoters are often differentially regulated. Studies by Benyajati *et al* (1983) have shown that alcohol dehydrogenase mRNA (from *D.Melanogaster*) differs in its 5' end during different developmental stages. Many studies have shown that PrP is developmentally regulated (Lieberburg *et al.*, 1987; McKinley *et al.*, 1987; Lazarini *et al.*, 1991 and Manson *et al.*, 1992a) but the process controlling this regulation are, as yet, not understood.

### 7.3 The use of transgenic mice to study PrP gene expression.

The use of tissue culture cell lines in the analysis of gene expression is often invaluable. Many facets of gene expression can be studied in this way. These often include mapping promoter elements and other regulatory elements such as enhancers and repressors. This is usually done by determining the effects of deletions on such elements. Studies involving tissue-specific expression of genes can also be performed in tissue culture cell lines. Tissue specific elements have been found in the rat insulin gene (Edlund *et al.*, 1985) and in the chicken δ-crystalline gene (Hayashi *et al.*, 1985) in this way. Both of these studies, incidentally, showed that tissue specific expression was mediated by elements found in the 5' flanking regions of the respective genes. It may be that elements which control tissue specific expression of PrP are also located in the 5' flanking region.

Despite the widespread use of cell culture for analysis of gene expression, it is often used for preliminary investigations. Although much valuable data can be obtained from tissue culture work, it is impossible to analyse sequences responsible for the correct temporal and spatial gene expression in the living animal. To do this one would have to produce transgenic mice. These mice contain either new or altered genes and these genes are present in every cell of the mouse. The most common way of creating a transgenic mouse is by direct microinjection of DNA into one of the pronuclei of a fertilised egg (for review see Palmiter and Brinster, 1986). To examine the expression of PrP one could use the PrP CAT constructs which were described in chapter 5. Although CAT has been used as a reporter gene to generate transgenic animals in the past (Julien *et al.*, 1990), both β-galactosidase and luciferase are generally preferred for such studies. This is because these proteins can be detected *in situ* rather than having to produce cell extracts for reporter protein measurement (for review see Alam and Cook, 1990). This makes both developmental and tissue specific gene expression easier to study. The objective of such experiments would be to determine the DNA sequences which control developmental and tissue specific expression of PrP. Would, for example, pCAT PrP 3.5 (described in chapter 5, but substituting CAT for lacZ) confer the normal pattern of PrP expression onto the reporter gene? For example, would lacZ be highly expressed in the brain, at intermediate levels in the heart and at low levels in the spleen (Oesch *et al.*, 1985; Caughey *et al.*, 1988)? Would lacZ be expressed at embryonic day 13.5 in the brain and spinal cord and then rise approximately 4-fold to adult day 20 (Manson *et al.*, 1991; Lazarini *et al.*, 1987)? If this was the case then one could conclude that all the sequences responsible for correct tissue specific and developmental expression were present within 1.2 kb of 5' flanking sequence and the first intron. If, however, abnormal expression of the reporter gene was observed then one would have to examine the role of the large intron (intron 2), 3' flanking region or more distal 5' flanking sequences.

## 7.4 PrP and Sinc: Relevance of flanking regions.

In mice, the major gene which controls scrapie incubation period is called *Sinc* (also known as Prn-i), of which there are two alleles: s7 and p7 (Dickinson *et al.*, 1968). Tight genetic linkage has been demonstrated between PrP and *Sinc* by crosses between s7 mice (NZW) and p7 mice (I/LnJ) as shown by Carlson *et al* (1986). Evidence that PrP and *Sinc* are congruent is given by restriction fragment length polymorphism (RFLP) analysis (Hunter *et al.*, 1987) and by direct nucleotide sequence analysis of mice with different *Sinc* genotypes (Westaway *et al.*, 1987). Sequence analysis of s7 mice predicts that they have a leucine residue at codon 108 and a threonine residue at codon 189. *Sinc* p7 mice have phenylalanine and valine residues at these respective codons.

Transgenetic studies have also been used in an attempt to define the relationship between PrP and *Sinc*. Westaway *et al.*, (1991) produced transgenic mice from a NZW background (*Sinc* s7 or Prn-a) which contained cosmid transgenes from an I/LnJ mouse (*Sinc* p7 or Prn-b). These transgenes were expected to confer a prolonged scrapie incubation period upon the transgenic mice when inoculated with scrapie agent. In fact, the opposite happened; mice carrying the *Sinc* p7 allele actually had shorter scrapie incubation periods than their s7 (NZW) counterparts. This occurred in four separate lines. This result was unexpected and several suggestions were proposed to explain it, including elevated levels of $PrP^c$ expression causing reduced incubation periods or aberrant transgene expression.

Alternatively, however it may be that the gene controlling incubation period lies adjacent to the PrP gene or is only partially represented within the genomic clone. Therefore, it may be that *Sinc* is very close to the PrP locus but that the two are not congruent. An examination of the flanking regions of the PrP gene may prove fruitful in proving if *Sinc* is distinct and separate from PrP. One way of doing this could be by looking for transcripts in the flanking regions of the PrP gene. Having identified transcripts, one could clone the cDNA of the gene and analyse it more fully.

## 7.5 Sporadic CJD: How is it caused?

Creutzfeldt-Jakob disease (CJD) can be divided into three basic forms: Iatrogenic, familial and sporadic.

Iatrogenic cases of CJD are those which are caused by accidental transmission. These accidental transmissions usually occur during surgical procedures, especially those involving the brain or the eye. The contamination in these cases either originates from graft tissue taken from a CJD affected individual or from contaminated surgical instruments which had presumably been used on a CJD patient and improperly decontaminated. Another example of iatrogenic CJD, involving no surgery, is given by individuals, usually children, who received human growth hormone treatments either subcutaneously or intramuscularly. Some of these patient subsequently developed CJD. A likely explanation for this is that pooled pituitary glands, from which the growth hormone was extracted, contained infectious agent. These pituitary glands had been inadvertently taken from CJD affected individuals (see Brown, 1988, for review).

Familial forms of CJD are those which are characterised by an autosomal form of inheritance (see Hsiao and Prusiner, 1990, for review). Familial forms of CJD are linked to mutations in the ORF of the PrP gene. A number of both point and insertion mutations have been described in this region (see section 1.4.2).

By far, the majority of CJD cases ( > 80 %) are of a sporadic (or "spontaneous") nature. The title of "sporadic" is generally attributed to a CJD case if there is no evidence of iatrogenic transmission or if no mutation in the ORF of PrP can be identified. A crucial point, however, is that much of PrP molecular biology is concentrated on the open reading frame exon (exon 2 in humans) as this is the exon which codes entirely for the protein. As a result of this no-one has looked at the 5' flanking exon or other possible regulatory regions of the gene for the existence of mutations. The possibility exists, therefore, that the incidence of sporadic CJD can be explained by mutations in the upstream exon of the PrP gene or possibly in regulatory regions such as the promoter or the intron. As exon 1 of the human PrP

gene (Puckett *et al.*, 1991) is only ~ 136 bp in length it would be relatively simple to determine if CJD patients harboured mutations in this region, by comparing nucleotide sequence to unaffected individuals. Data presented in chapter 5 is the first evidence of regulatory elements being found in the 5' flanking region of a PrP gene in any species. Examination of the equivalent region of the human PrP gene for mutations could also be performed. In the absence of data relating to the existence of other regulatory elements occurring elsewhere in the gene, it would be difficult to determine if mutations in such regions are linked to sporadic CJD.

Sporadic CJD may be explained, however, in other ways such as by a rare infection of a ubiquitous agent or by development of an age related somatic mutation occurring somewhere on the PrP gene (Hsiao and Prusiner, 1990). In the absence of experimental data, however, the possibility of mutations occurring in exon 1 or, in the promoter region, and causing sporadic CJD, still exists.

### 7.6 Is PrP a housekeeping gene ?

Although there are exceptions, promoter regions of genes which do not contain either a TATA or CAAT box are generally associated with a housekeeping function. These housekeeping genes are expressed in many, if not all, cell types and often perform essential metabolic function e.g dihydrofolate reductase (Mitchell *et al.*, 1986) and adenosine deaminase (Valerio *et al.*, 1985). The promoter regions of housekeeping genes generally share common features which make them distinct from promoters of other genes. These include the absence of both TATA and CAAT boxes, highly G C rich, multiple transcription start sites, and multiple binding sites for the transcription factor Sp1 (see Dynan, 1986, for review). Data presented in chapters 4 and 6 show that the mouse PrP promoter also shares these features.

Although the normal function of PrP is not known, despite the generation of PrP null mice (Bueler *et al.*, 1992; Manson *et al.*, 1994), its widespread expression has been well documented (Oesch *et al.*, 1985; Caughey *et al.*, 1988). PrP gene expression was shown in these studies to be high in the brain but expressed very poorly in the spleen. Intermediate levels of PrP mRNA were found in organs such as the heart and lung. These findings imply that PrP gene expression can be regulated. In chapter 4 it was shown that a CpG island existed in the 5' region of the PrP gene. It may be that differential methylation of this CpG island controls the level of PrP gene expression in different tissues. Most other housekeeping genes have CpG islands associated at their 5' end (Gardiner-Garden and Frommer, 1987), and a similar degree of CpG island methylation may result in similar levels of gene expression in different tissues.

Regulated gene expression of housekeeping genes such as dihydrofolate reductase (dhfr) and 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG coA reductase) has been shown to occur. For example, HMG coA reductase is a housekeeping enzyme which is involved in cholesterol biosynthesis. A negative feedback mechanism exists whereby cholesterol can suppress HMG coA reductase at the level of transcription (Luskey *et al.*, 1983). Dhfr, on the other hand, has been shown to have its rate of transcription increased four-fold during the proliferative growth phase of fibroblasts (Hendrickson *et al.*, 1980; Santiago *et al.*, 1983).

It therefore seems likely that PrP is a housekeeping gene and, in common with other housekeeping genes, its expression may be regulated under certain conditions.

## 7.7 Understanding PrP gene regulation: Implications for therapeutic intervention.

Data now exists which show that the levels of normal PrP ($PrP^c$) are critical in controlling aspects of disease such as scrapie incubation period and progression of disease.

Two main experiments have provided these data. Firstly, Scott *et al* (1989) generated mice which were transgenic for the Syrian hamster PrP gene (see section 1.6). Various lines had copy numbers for the transgene ranging from 2 to ~ 60 copies. After challenging these mice with hamster scrapie agent it was found that mice with high transgene copy number, and subsequently high $PrP^c$ and PrP mRNA expression (Prusiner *et al.*, 1990), had much shorter incubation periods than mice with lower transgene copy numbers. The idea that PrP expression and scrapie incubation period inversely correlate was later reinforced by Bueler *et al* (1992) who generated PrP homozygous null ($PrP^{0/0}$) and heterozygous null ($PrP^{0/+}$) mice. These mice were challenged with scrapie agent and their incubation periods ascertained (Bueler *et al.*, 1993). Homozygous null mice were found to be completely resistant to scrapie up to the age of 14 months. Wild-type mice all developed disease and died within 180 days. Heterozygous null mice showed an enhanced resistance to scrapie, showing clinical symptoms after 253-337 days, but remaining alive after 380 days. In addition, the progression of disease (i.e clinical onset to death) is much slower in heterozygous null mice (50 - 130 days) than in wild type mice (~ 13 days).

A possible consequence of understanding PrP gene regulation is that development or progression of the transmissible degenerative encephalopathies may be controlled. Therapeutic intervention, by such means, would be less drastic than ablating PrP genes. This intervention may conceivably come about in several ways. If one, for example, had a thorough understanding of the factors which enhanced and repressed expression of PrP, especially in the central nervous system, then one may be able to control the expression of PrP in a given tissue. Consequently, the levels of expressed

PrP may be lowered beneath a critical level whereby disease could not be manifested within the life span of the animal.

Another therapeutic approach has been suggested by Weissman *et al* (1993), whereby the levels of PrP expression could be controlled by antisense oligonucleotide therapy. Much progress has been made in the treatment of diseases with a genetic predisposition such as cystic fibrosis. For example, Hyde *et al* (1993) have introduced functional copies of the cystic fibrosis transmembrane conductance regulator (CTFR) gene into the airways of transgenic mice which carry defective copies of the gene. These functional copies of the gene were introduced by aerosol mediated lipofection and were found to induce correction of the ion conductance defect. This, in effect, caused transient alleviation of disease symptoms.

However, in the TDEs the main focus of disease is in the central nervous system and therefore targeting of potential treatments may be more difficult. Despite this, it seems that if a therapeutic approach to the TDEs were to be developed then an understanding of PrP gene expression would be crucial.

**Alam, J. and Cook, J.C.** (1990). Reporter genes: Application to the study of mammalian gene transcription. *Anal. Biochem.* **188**, 245-254.

**Alper, T., Cramp, W.A., Haig, D.A. and Clarke, M.C.** (1967). Does the scrapie agent replicate without nucleic acids? *Nature* **214**, 764-766.

**Alper, T., Haig, D.A., and Clarke, M.C.** (1978). The scrapie agent: Evidence against its dependence for replication on intrinsic nucleic acid. *J. Gen. Virol.* **41**, 503-516.

**Angel, P., Imagawa, M., Chlu, M., Stein, B., Imbra, R.J., Rahmsdorf, H.J, Jonat, C., Herrlich P., and Karin, M.** (1987). Phorbol ester-inducible genes contain a common cis element recognised by TPA-modulated transacting factor. *Cell* **49**, 729-739.

**Banerji, J., Rusconi, S. and Schaffner, W.** (1981). Expression of a globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299-308.

**Basler, K., Oesch, B., Scott, M., Westaway, D., Walchi, M., Groth, D.F., McKinley, M.P., Prusiner, S.B., and Weissmann, C.** (1986). Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene. *Cell* **46**, 417-428.

**Bellinger-Kawahara, C., Cleaver, J.E., Diener, T.O., and Prusiner, S.B.** (1987). Purified scrapie prions resist inactivation by U.V irradiation. *Virology* **160**, 271-274.

**Bellinger-Kawahara, C., Kempner, E., Groth, D., Gabizon, R. and Prusiner, S.B.** (1988). Scrapie prion liposomes and rods exhibit target sizes of 55,000 Da. *Virology* **165**, 537-541.

**Bendheim, D.C., Potemska, A., Kascsak, R., and Bolton, D.C.** (1988). Purification and partial characterisation of the cellular homologue of the scrapie agent protein. *J. Inf. Dis.* **158**, 1198-1208.

**Benoist, C. and Chambon, P.** (1980). In vivo sequence requirements of the SV40 early promoter region. *Nature* **290**, 304-310.

**Benyajati, C., Spoerei, N., Haymerle, H., and Ashburner, M.** (1983). The messenger RNA for alcohol dehydrogenase in Drosophila melanogaster differs in its 5' end in different developmental stages. *Cell* **33**, 125-133.

**Bergers, G., Reikerstorfer, A., Braselmann, S., Graninger, P., and Busslinger, M.** (1994). Alternative promoter usage of the Fos responsive gene *Fit-1* generates mRNA isoforms coding for either secreted or membrane bound proteins related to the Il-1receptor. *EMBO. J.* **13**(5), 1176-1188.

**Bird, A.P.** (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209-213.

**Birnboim, H.C., and Doly, J.** (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.

**Bolton, D.C., McKinley, M.P., and Prusiner, S.B.** (1982). Identification of a protein that purifies with the scrapie prion. *Science* **218**, 1309-1311.

**Bolton, D.C., Meyer, R.K., and Prusiner, S.B.** (1985). Scrapie PrP 27-30 is a sialoglycoprotein. *J. Virol.* **53**, 596-606.

**Bolton, D.C., Bendheim, P.E., Marmorstein, A.D., Potemska, A.** (1987). Isolation and structural studies of the intact scrapie agent protein. *Arch. Biochem. Biophys.* **258**, 579-590.

**Borchelt, D.R., Scott, M., Taraboulos, A., Stahl, N., Prusiner, S.B.** (1990). Scrapie and cellular prion proteins differ in their kinetics of synthesis and topology in cultured cells. *J. Cell. Biol.* **110**, 743-752.

**Bornstein, B., and McKay, J.** (1988). The first intron of the alpha 1 (I) collagen gene contains several transcriptional regulatory elements. *J. Biol. Chem.* **263** (4), 1603-1606.

**Breathnach, R., and Chambon, P.** (1981). Organisation and expression of eukaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**, 349-383.

**Brown P.** 1988. The clinical neurology and epidemiology of Creutzfeldt-Jakob disease, with special reference to iatrogenic cases. *In Bock, G., and Marsh, J.* (Eds). Novel infectious agents and the central nervous system. Wiley, Chichester (Ciba Found Symposium, **135**). pp 3-18.

**Brown, H.R., Goller, N.L., Rudelli, R.D., Merz, G.S., Wolfe, G.C., Wisniewski, H.M., and Robakis, N.K.** (1990). The mRNA encoding the scrapie agent is present in a variety of non-neuronal cells. *Acta. Neuropath.* **80**, 1-6.

**Brown, P., Goldfarb, L.G., Kovanen, J., Haltia, M., Cathala, F., Sulima M., Gibbs, C.J., Gajdusek, D.C.** (1992). Phenotypic characteristics of familial Creutzfeldt-Jakob disease associated with codon 178 Asn PRNP mutation. *Ann. Neurol.* **31**, 282-285.

**Bruce, M.E., and Dickinson, A.G.** (1987). Biological evidence that scrapie agent has an independent genome. *J. Gen. Virol.* **68**, 79-89.

**Bruce, M.E., McBride, P.A., and Farquhar, C.F.** (1989). Precise targeting of the pathology of the sialoglycoprotein, PrP, and vacuolar degeneration in mouse scrapie. *Neurosci. Lett.* **102**, 1-6.

**Bruce, M.E., McConnel, I., Fraser, H., and Dickinson, A.G.** (1991). The disease characteristics of different strains of scrapie in Sinc congenic mouse lines: Implications for the nature of the agent and host control of pathogenesis. *J. Gen. Virol.* **72**, 595-603.

**Bueler, H., Fischer, M., Lang, Y., Bluethman, H., Lipp, H.P., DeArmond, S.J., Prusiner, S.B., Aguet, M., and Weissmann, C.** (1992). Normal development and behavour of mice lacking the neuronal cell-surface PrP protein. *Nature* **356**, 577-582.

**Bueler, H., Aguzzi, A., Sailer, A., Greiner, R.A., Autenried, P., Aguet, M.** (1993). Mice devoid of PrP are resistant to scrapie. *Cell* **73**, 1339-1347.

**Burke, D.T., Carle, G.F. and Olson, M.V.** (1987). Cloning of large segments of exogenous DNA into yeast by means of artifical chromosome. *Science* **236**, 806-817.

**Butler, D.A., Scott, M.R.D., Bockman, J.M., Borchelt, D.R., Taraboulos, A., Hsiao, K., Kingsbury, D.T., and Prusiner, S.B.** (1988). Scrapie infected murine neuroblastoma cells produce protease resistant prion proteins. *J. Virol.* **62**, 1558-1564.

**Cao, X., Kosk, R.A., Gashler, A., McKiernan, M., Morris, C.F., Gaffney, R., Hay, R.V. and Sukhatme, V.P.** (1990). Identification and characterisation of Egr-1 gene product, a DNA-binding zinc finger protein induced by differentiation and growth signals. *Mol. Cell. Biol.* **10**, 1931-1939.

**Carlson, G. A., Kingsbury, D.T., Goodman, P.A., Coleman, S., Marshall, S.T., DeArmond, S., Westaway, D., and Prusiner, S.B.** (1986). Linkage of prion protein and scrapie incubation time genes. *Cell* **46**, 503-511.

**Cashman, N.R., Loertschner, R., Nalbantoglu, J., Shaw, I., Kascsak, R.J., Bolton, D.C., Bendheim, P.E.** (1990). Cellular isoform of the scrapie agent protein participates in lymphocyte activation. *Cell* **61**, 185-192.

**Caughey, B., Race, R.E., and Chesebro, B.** (1988). Detection of prion protein mRNA in normal and scrapie infected tissues and cell lines. *J. Gen. Virol.* **69**, 711-716.

**Chen, C., and Okayama, H.** (1987). High efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, (8) 2745-2757.

**Chesebro, B., Race, R., Wehrly, K., Nishio, J., Bloom, M., Lechner, D., Bergstrom, S., Robbins, K., Mayer, L., Keith, J., Garon, C., and Haase, A.** (1985). Identification of scrapie prion protein specific mRNA in scrapie infected and uninfected brain. *Nature* **315**, 331-333.

**Clarke, L., and Carbon, J.** (1976). A colony bank containing synthetic Col EI hybrid plasmids representative of the entire E.coli genome. *Cell* **9**, 91-99

**Clinton, J., Forsyth, C., Royston, M.C. and Roberts, G.W.** (1993). Synaptic neurodegeneration is the primary neurological feature in prion disease: a preliminary study. *Neuroreport* **4**, 65-68.

**Collinge, J., Whittington, M.A., Sidle, K.C.L., Smith, C.J., Palmer, M.S., Clarke, A.R., and Jeffreys, J.G.R.** (1994). Prion protein is necessary for normal synaptic function. *Nature* **370**, 295-297.

**Coulondre, C., Miller, J.H., Farabaugh, P.J., and Gilbert, W.** (1978). Molecular basis of base substituition hotspots in *Escherichia coli*. *Nature* **274**, 775-780.

**De Wet, J.R., Wood, K.V., De Luca, M., Helinski, D.R., Subramani, S.** (1987). Firefly luciferase gene: structure and expression in mammalian cells. *Mol.Cell. Biol.* 7(2) , 725-737.

**Devereux, J., Haeberli, P., and Smithies, O.** (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12** (1), 387-395.

**Dickinson, A.G., Meikle, V.M.H., and Fraser, H.** (1968). Identification of a gene which controls the incubation period of some strains of scrapie agent in mice. *J. Comp. Pathol.* **78**, 293-299.

**Dickinson, A.G., Stamp, J.T., Renwick, C.C.** (1974). Maternal and lateral transmission of scrapie in sheep. *J. Comp. Path.* **84**, 19-25.

**Dickinson, A.G, and Fraser, H.** 1977. Scrapie pathogenesis in inbred mice: an assessment of host control and response involving many strains of agent. *In: Termeulen, V., and Katz, M.* (Eds). Slow virus infections of the central nervous system. Springer, Berlin Heidelberg New York. pp 3-14.

**Dlouhy, S.R., Hsiao, K., Farlow, M.R., Faroud, T., Conneally, P.M., Johnson, P.M., Prusiner, S.B., Hodes, M.E., and Ghetti, B.** (1992). Linkage of the Indiana kindred of Gerstmann-Straussler-Scheinker disease to the prion protein gene. *Nature Genetics* **1**, 64-67.

**Doh-ura, K., Tateishi, J., Sasaki, H., Kimamato, T., and Sakaki, Y.** (1989). Pro→ Leu change at position 102 of prion protein is the most common but not the sole mutation related to Gerstmann-Straussler syndrome. *Biochem. Biophys. res. comm.* **163** (2), 974-979.

**Donehower, L.A., Harvey, M., Slage, B.L., McArthur, M.J., Montgomery C.A., Butel, J.S., and Bradley, A.** (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumors. *Nature* **356**, 215-221.

**Donis, J. A., Ventosa-Michelman, M., and Neve, R.L.** (1993). Comparison of expression of a series of mammalian vector promoters in the neuronal cell lines PC12 and HT4. *Biotechniques* **15**, 786-787.

**Duguid, J.R., and Dinaver, M.C.** (1990). Library subtraction of *in vitro* cDNA libraries to identify differentially expressed genes in scrapie infection. *Nucleic Acids Research* **18** (9) 2789-2792.

**Duguid, J.R., and Trezepaz, C.** (1993). Major histocompatibiliy complex genes have an increased brain expression after scrapie infection. *Proc. Natl. Acad. Sci. (USA).* **90**, 114-117.

**Dynan, W.S., and Tjian, R.** (1983). The promoter specific transcription factor Sp1binds to upstream sequences in the SV40 promoter. *Cell* **35**, 79-87.

**Dynan, W.S.** (1986). Promters for housekeeping genes. *Trends in genetics* **2**, August. 196-197.

**Ehrich, E., Craig, A., Poustka, A., Frischauf, A.M., Lehrach, H.** (1987). A family of cosmid vectors with the multi-copy R6K replication origin. *Gene* **57**, 229-237.

**Edlund, T., Walker, M.D., Barr, P.J., and Rutter, W.J.** (1985). Cell specific expression of the rat insulin gene: Evidence for the role of two distinct 5' flanking elements. *Science* **230**, 912-916.

**Eklund, C.M., Kennedy, R.L., Hadlow, W.J.** (1967). Pathogenesis of scrapie virus infection in the mouse. *J. Inf. Disease.* **131**, 104-110.

**Faisst, S., and Meyer, S.** (1992). Compilation of vertebrate encoded transcription factors. *Nucleic Acids Research* **20**, (1) 3-26.

**Feinberg, A.P., and Vogelstein, B.** (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specifc activity. *Anal. Biochem.* **132**, 6-13.

**Felgner, D.C., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M, Northrop, J.P., Ringold, G., and Danielsen, M.** (1987). Lipofection: A highly efficient, lipid mediated, DNA transfection procedure. *Proc. Natl. Acad. Sci.(USA).* **84**, 7413-7417.

**Gajdusek, D.C.** (1977). Unconventional viruses and the origin and disappearance of kuru. *Science* **197**, 943-960.

**Gardiner-Garden, M., and Frommer, M.** (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261-282.

**Garfinkel, L.I. and Davidson, N.** (1987). Developmentally regulated expression of a truncated myosin light chain $1_F/3_F$ gene. *Mol. Cell. Biol.* 7(10), 3826-3829.

**Ghosh, D.** (1990). A relational database of transcription factors. *Nucleic Acids Research.* **18** (7), 1749-1756.

**Gibson, T.J., Coulson, A.R., Sulston, J.E., Little, P.F.R.** (1987). Lorist2, a cosmid with transcriptional terminators insulating vector genes from interference by promoters within the insert: effect on DNA yield and cloned insert frequency. *Gene* **53**, 275-281.

**Goldfarb, L.G., Mitrova, E., Brown, P., Toh, B.H., Gajdusek, D.C.** (1990) a. Mutation in codon 200 of scrapie amyloid protein gene in two clusters of Creutzfeldt-Jakob disease in Slovakia. *Lancet* **336**, 514-515.

**Goldfarb, L.G., Korczyn, A., Brown, P., Chapman, J., Gajdusek, D.C.** (1990) b. Mutation in codon 200 of scrapie amyloid precursor gene linked to Creutzfeldt-Jakob disease in sephardic Jews of Libyan and non-Libyan origin. *Lancet* **336**, 637-638.

**Goldfarb, L.G., Brown, P., McCombie, W.R., Goldgaber, D., Swingold, G.D, Wills, P., Cervenakova, L., Baron, H., Gibbs, C.J., and Gajdusek, D.C.** (1991). Transmissible familial Creutzfeldt-Jakob disease associated with five, seven and eight octapeptide coding repeats in the PRNP gene. *Proc. Natl. Acad. Sci. (USA).* **88**, 10926-10930.

**Goldmann, W., Hunter, N., Foster, J.D., Salbaum, J.M., Beyreuther, K, Hope, J.** (1990). Two alleles of a neural protein gene linked to scrapie in sheep. *Proc. Natl. Acad. Sci.* (USA). **87**, 2476-2480.

**Goldmann, W.** (1993). PrP gene and its association with spongiform encephalopathies. *British Medical Bulletin* **49** (4), 839-859.

**Goosen, B., Caughman, S.W., Harford, J.B., Klausner, R.D., Hentze, M.W.** (1990). Translational repression by a complex between the iron-responsive element of ferritin mRNA and its specific cytoplasmic binding protein is position-dependent *in vivo*. *EMBO J.* **9**, 4127-4133.

**Graham, F.L. and van Der Eb, A.J.** (1973). A new technique for assay of infectivity of human adenovirus 5 DNA. *Virology*. **52**, 456-467.

**Grunstein, M., and Hogness, D.S.** (1975). Colony hybridisation: A method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. (USA)*. **72**, 3961-3965.

**Hayashi, S., Kondoh, H., Yasuda, K., Gen-ichiro, S., and Okada, T.S.** (1985). Tissue-specific regulation of a chicken δ-crystallin gene in mouse cells: involvement of the 5' end region. *EMBO. J.* **4** (9), 2201-2207.

**Heinikoff, S.** (1984). Unidirectional deletions with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351-359.

**Hendrickson, S.L., Wu, J.R., and Johnson, L.F.** (1980). Cell cycle regulation of dihydrofolate reductase mRNA in mouse fibroblasts. *Proc. Natl. Acad. Sci. (USA)*. **77**, 5140-5144.

**Herrick, K.R., Gorin, F.A., Park, E.A., Tart, R.C.** (1993). Characterisation of the 5' flanking region of the gene encoding rat liver glycogen phosphorylase. *Gene*. **126**, 203-211.

**Higgs, D.R., Wood, W.G., Jarman, A.P., Sharpe, J., Lida, J., Pretorius, I.M., Ayyub, H.** (1990). A major positive regualatory region located far upstream of the human α-globin gene locus. *Genes and Development* **4**, 1588-1601.

**Hope, J., Morton, L.J.D., Farquar, C.F., Multhaup, G., Beyreuther, K., Kimberlin, R.H.** (1986). The major polypeptide of scrapie associated fibrils (SAF) has the same size, charge distribution and N-terminal protein sequence as predicted for the normal brain protein (PrP). *EMBO J.* **5**, 2591-2597.

**Hope, J., and Manson, J.** (1991). The scrapie fibril protein and its cellular isoform. *Curr. top. micr. immunol* **173**, 57-74.

**Hsiao, K., Baker, H.F., Crow, T.J., Poulter, M., Owen, F., Terwilliger, J.D., Westaway, D., Ott, J. and Prusiner, S.B.** (1990). Linkage of a prion protein missense variant to Gerstmann-Straussler syndrome. *Nature* **338**, 342-345.

**Hsiao, K., and Prusiner, SB.** (1990). Inherited human prion diseases. *Neurology* **40**, 1820-1827.

**Hsiao, K., Dlouhy, S.R., Farlow, M.R., Cass, C., Dacosta, M., Conneally, P.M., Hodes, M.E., Ghetti, B., and Prusiner, S.B.** (1992). Mutant prion proteins in Gerstmann-Straussler-Scheinker disease with neurofibrillary tangles. *Nature Genetics* **1**, 68-71.

**Hunter, N., Hope, J., Mc Connel, I., and Dickinson, A.G.** (1987). Linkage of the scrapie associated fibril protein (PrP) gene and Sinc using congenic mice and restriction fragment length polymorphism analysis. *J. Gen. Virol.* **68**, 2711-2716.

**Hunter, N., Foster, J.D., Dickinson, A.G. and Hope, J.** (1989). Linkage of the gene for scrapie associated fibril protein (PrP) to the Sip gene in Cheviot sheep. *Vet.Rec.* **124**, 364-366.

**Hyde, S.C., Gill, D.R., Higgins, C.F., Trezise, H.E.O., MacVinish, L.J., Cuthbert, A.W., Ratcliff, R., Evans, M.J., and Colledge, W.T.** (1993). Correction of the ion transport defect in cystic fibrosis transgenic mice by gene therapy. *Nature* **362**, 250-255.

**Imagawa, M., Chiu, R., and Karin, M.** (1987). Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase and cAMP. *Cell* **51**, 251-260.

**Izumi, R., Yamada, T., Yoshikai, S., Sasaki, H., Hattori, M., Sakaki Y.** (1992). Positive and negative regulatory elements for the expression of the Alzheimer's disease amyloid precursor encoding gene in mouse. *Gene* **112**, 189-195.

**Josse, J., Kaiser, A.D. and Kornberg, A.** (1961). Frequencies of nearest neighbor base sequences in DNA. *J. Biol. Chem.* **236**, 864-875.

**Julien, J.P., Beaudet, C., Tretjakoff, I., and Peterson, A.** (1990). Neurofilament gene expression in transgenic mice. *J. Physiol.* **84**, 50-52.

**Kellings, K., Meyer, N., Mirenda, C., Prusiner, S.B., Riesner, D.** (1992). Further analysis of nucleic acids in purified scrapie prion preparations by improved return refocussing gel electrophoresis (RRGE). *J. Gen. Virol.* **73**, 1025-1029.

**Kingsbury, D.T., Kasper, K.C., Stites, D.D., Watson, J.D., Hogan, R.N., and Prusiner, S.B.** (1983). Genetic control of scrapie and Creuzfeldt-Jacob disease in mice. *J. Immunol.* **131**, 491-496.

**Klein, T.M., Fromm, M., Weissinger, A., Tomes, D., Schaaf, S., Sletten, M., and Sanford, J.C.** (1988). Transfer of foreign genes into intact maize cells with high-velocity microprojectiles. *Proc. Natl. Acad. Sci. (USA).* **85**, 4305-4309.

**Kretzschmar, H.A., Strowring, L.E., Westaway, D., Stubblebine, W.H., Prusiner, S.B., De Armond, S.J.** (1986). Molecular cloning of a human prion protein cDNA. *DNA.* **5**, 315-324.

**Kretzschmar, H.A., Prusiner, S.B., Strowring, L.E., De Armond, S.J.** (1986). Scrapie prion proteins are synthesised in neurones. *Am. J. Path.* **122**, 1-5.

**Kwon, Y.K. and Hecht, N.B.** (1993). Binding of a phosphoprotein to the 3' untranslated region of the mouse protamine 2 mRNA temporally represses its translation. *Mol.Cell.Biol.* **13** (10), 6547-6557.

**Lamb, N.J.C., Fernandez, A., Tourkine, N., Jeanteur, P., Blanchard, J.M.** (1990). Demonstration in living cells of an intragenic negative regulatory element within the rodent c-fos gene. *Cell* **61**, 485-496.

**Lazarini, F., Deslys, J.P., Dormont, D.** (1991). Regulation of glial fibrillary acidic protein, actin and prion protein mRNAs during brain development in mouse. *Molec. Brain. Res.* **10**, 343-346.

**Lemaire, P., Vesque, C., Schmitt, J., Stunnenberg, H., Reiner, F., and Charnay, P.** (1990). The serum inducible mouse gene Krox-24 encodes a sequence specific transcriptional activator. *Mol. Cell. Biol.* **10**, 3456-3467.

**Liao, Y.J., Lebo, R.V., Clawson, G.A., Smuckler, E.A.** (1986). Human prion protein cDNA: Molecular cloning, chromosomal mapping and biological implications. *Science* **233**, 364-367.

**Lieberburg, I.** (1987). Developmental expression and regional distribution of the scrapie associated protein mRNA in the rat central nervous system. *Brain Res.* **417**, 363-366.

**Locht, C., Cheesbro, B., Race, R., Kerth, J.M.** (1986). Molecular cloning and complete sequence of prion protein cDNA from mouse brain infected with scrapie agent. Proc. *Natl. Acad. Sci. (USA).* **83**, 6372-6376.

**Lowenstein, D.H., Butler, D.A., Westaway, D., McKinley, M.P., DeArmond S.J., Prusiner, S.B.** (1990). Three hamster species with different scrapie incubation times and neuropathological features encode distinct prion proteins. *Mol. Cell. Biol.* **10**, 1153-1163.

**Luskey K.L., Faust, J.R., Chin, D.J., Brown, M.S., Goldstein, J.L..** (1983). Amplification of a gene for 3-hydroxy-3-methylglutaryl coenzyme A reductase, but not for the 53 kDa protein, in UT-1 cells. *J. Biol. Chem.* **258**, 8462-8469.

**Maizel, J.V., and Lenk, R.P.** (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proc. Natl. Acad. Sci. (USA).* **78**, 7665-7669.

**Manetto, V., Medori, R., Cortelli, P., Montagna, P., Tinuper, P., Baruzzi, A., Rancurel, G., Hauw, J.J., Vanderhaeghen, J.J., Mailleux, P., Bugiani, O., Tagliavini, F., Bouras, C., Rizzuto, N., Lugaresi, E., and Gambetti, P.** (1992). Fatal familial insomnia: clinical and pathological study of five new cases. *Neurology* **42**, 312-319.

**Manson, J., West, J.D., Thomson, V., McBride, P., Kaufman, M.H., Hope, J.** (1992)a. The prion protein gene: a role in mouse embryogenesis? *Development* **115**, 117-122.

**Manson, J., McBride, P., Hope, J.** (1992) b. Expression of the PrP gene in the brain of *Sinc* congenic mice and its relationship to the development of scrapie. *Neurodegeneration.* **1**, 45-52.

**Manson, J.C., Clarke, A.R., Hooper, M.L., Aitchison, L., McConnell, I., and Hope, J.** (1994). 129/Ola mice carrying a null mutation in PrP that abolishes mRNA production are developmentally normal. *Molecular Neurobiology* **8**, 121-127.

**Masters, C.L., Gajdusek, D.C., Gibbs, C.J.** (1981). The familial occurence of Creutzfeldt-Jakob disease and Alzheimers disease. *Brain* **104**, 535-558.

**Masters, C.L., Simms, G., Weiman, N.A., Multhaup, G., McDonald, B.L., Beyreuther K.** (1985). Amyloid plaque core protein in Alzheimers disease and Downs syndrome. *Proc. Natl. Acad. Sci. (USA).* **82**, 4245-4249.

**McCutchan, J.H., and Pagano, J.** (1968). Enhancement of the infectivity of Simian virus 40 deoxyribonucleic acid with diethyl-aminoethyl-D-dextran. J. Natl. cancer Inst. **41** (2), 351-357.

**McKinley, M.P., Bolton, D.C., and Prusiner, S.B.** (1983). A protease resistant protein is a structural component of the scrapie prion. *Cell* **35**, 57-62.

**McKinley, M.P., Hay, B., Lingappa, V.R., Lieberburg, I., and Prusiner, S.B.** (1987). Developmental expression of prion protein gene in brain. *Developmental biology* **121**, 105-110.

**McKinley, M.P., Taraboulos, A., and Kenega, L.** (1991). Ultrastructural localisation of scrapie prion proteins in cytoplasmic vesicles of infected cells. *Lab. Invest.* **65**, 622-630.

**Medori, R., Tritschier, H-J., LeBlanc, A., Villare, F., Manetto, V., Chen, H.Y., Xue, R., Leal, S., Montagna, P., Cortelli, P., Tinuper, P., Avoni P., Mochi, M., Baruzzi, A., Hauw, J.J., Ott, J., Lugaresi, E., Gambetti, L., and Gambetti, P.** (1992). Fatal familial insomnia, a prion disease with a mutation at codon 178 of the prion protein gene. *New. England. J. Med.* **326**, 444-449.

**Meyer, R.K., McKinley, M.P., Bowman, K.A., Braunfeld, M.B., Barry, R.A., Prusiner, S.B.** (1986). Separation and properties of cellular and scrapie prion proteins. *Proc. Natl. Acad. Sci. (USA).* **83**, 2310-2314.

**Meyer, N., Rosenbaum, V., Schmidt, B., Gilles, K., Mirenda, C., Groth, D., Prusiner, S.B., and Riesne,r D.** (1991). Search for a putative scrapie enome in purified scrapie fractions reveals a paucity of nucleic acids. *J.Gen. Virol.* **72**, 37-49.

**Minowa, M.T., Minowa, T., Monsma, F.J., Sibley, D.R., Mouradian, M.** (1992). Characterisation of the 5' flanking region of the human D1a dopamine receptor. *Proc. Natl. Acad. Sci. (USA).* **89**, 3045-3049.

**Mitchell, P.J., Carothers, A.M., Jang, H.H., Harding, J.D., Kas, E., Nenolia, L., and Chasin, L.A.** (1986). Multiple transcription start sites, DNaseI hypersensitive sites, and an opposite strand exon in the 5' region of the CHO *dhfr* gene. *Mol. Cell. Biol.* **6** (2), 425-440.

**Mitchell, P.J., Wang, C., and Tjian, R.** (1987). Positive and negative regulation of transcription *in vitro*: enhancer binding protein AP2 is inhibited by SV40 T-antigen. *Cell* **56**, 847-861.

**Mobley, W.C., Neve, R.L., Prusiner, S.B., McKinley, M.P.** (1988). Nerve growth factor increases mRNA levels for the prion protein and β-amyloid protein precursor in developing hamster brain. *Proc. Natl. Acad. Sci. (USA).* **85**, 9811-9815.

**Morris, T., Marashi, F., Weber, L., Hickey, E., Greenspan, D., Bonner, J., Stein, J., and Stein, G.** (1986). Involvement of the 5' leader sequence in coupling the stability of a human H3 histone mRNA with DNA replication. *Proc. Natl. Acad. Sci. (USA).* **83**, 981-985.

**Mount, SM.** (1982). A catalogue of splice junction sequences. *Nucleic Acids Research.* **10**, 459-472.

**Nomenclature commitee of the international union of biochemistry.** (1985). Nomenclature for incompletely specified bases in nucleic acid sequences. *Eur. J. Biochem.* **150**, 1-5.

**Oesch, B., Westaway, D., Walchi, M., McKinley, M.P., Kent, S.B.H., Aebersold, R., Barry, R.A., Tempst, P., Teplow, D.B., Hood, L.E., Prusiner, S.B., and Weissmann, C.** (1985). A cellular gene encodes scrapie PrP 27-30 protein. *Cell.* **40**, 735-746.

**Olescak, E.L., Murdoch, G., Manuelidis, L., Manuelidis, E.E.** (1989). Growth factor production by Creutzfeldt-Jakob disease cell lines. *J. Virol.* **62**, 3103-3108.

**Outram, G.W.** (1976). The pathogenesis of scrapie in mice. *In : Kimberlin, R.H.* (Ed). Slow virus diseases of animals and man. North-Holland, Amsterdam pp325-357.

**Owen, F., Poulter, M., Collinge, J., Leach, M., Shah, T., Lofthouse, R., Chen, Y., Crow, T.J., Harding, A.E., Hardy, J., and Rossor, M.N.** (1991). Insertions in the prion protein gene in atypical dementias. *Experimental neurology* **112**, 240-242.

**Palmiter, R.D., and Brinster, R.L.** (1986). Germ line transformation of mice. *Ann. Rev. Genet.* **20**, (4) 65-99.

**Pelletier, J., and Sonenberg, N.** (1985). Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**, 515-526.

**Piechaczyk, M., Yang, J.Q., Blanchard, J.M., Jeanteur, P., and Marcu, K.B.** (1985). Post-transcriptional mechanisms are responsible for accumulation of truncated c-myc RNAs in murine plasma cell tumours. *Cell* **42**, 589-597.

**Potter, H., Weir, L., Leder, P.** (1984). Enhancer dependent expression of human κ immunoglobulain genes introduced into pre-B lymphocytes by electroporation. *Proc. Natl. Acad. Sci. (USA).* **81**, 7161-7165.

**Prusiner, S.B.** (1982). Novel proteinaceous infectious particles cause scrapie. *Science* **216**, 136-144.

**Prusiner, S.B., Scott, M., Foster, D., Pan, K.M., Groth, D., Mirenda, C., Torchia, M., Yang, S.L., Serban, D., Carlson, G.A., Hoppe, P.C., Westaway, D., DeArmond, S.J.** (1990). Transgenetic studies implicate interactions between homologous PrP isoforms in scrapie prion replication. *Cell* **63**, 673-686.

**Puckett, C., Concannon, P., Casey. C., Hood, L.E.** (1991). Genomic structure of the human prion protein gene. *American Journal of Human Genetics*. **49**, 320-329.

**Rabbits P.H., Forster, H., Stinson, M.A., and Rabbits, T.H.** (1985). Truncation of exon 1 from the c-myc gene results in prolonged c-myc RNA stability. *EMBO J.* **4** (13), 3727-3733

**Race, R.E., Fadness, L.H., Chesebro, B.** (1987). Characterisation of scrapie infection in mouse neuroblastoma cells. *J. Gen. Virol.* **68**, 1391-1399.

**Race, R.E., Caughey, B., Graham, K., Ernst, D., Chesebro, B.** (1988). Analyses of the frequency of infection, specific infectivity, and prion protein biosynthesis in scrapie infected neuroblastoma cell lines. *J. Virol.* **67**, 7845-7849.

**Roberts, G.W., Lofthouse, R., Brown, R., Crow, T.J., Barry, P.A., Prusiner, S.B.** (1986). Prion protein immunoreactivity in human transmissible dementias. *New. England. J. Med.* **315**, 1231-1233.

**Rosenthal, N.** (1987). Identification of regulatory elements of cloned genes with functional assays. *Methods in Enzymology* **152**, 704-720.

**Rotheneder, H., Grabner, M., Wintersberger, E.** (1991). Presence of regulatory sequences within intron 2 of the mouse thymidine kinase gene. *Nucleic Acids Research* **19**, 6805-6809.

**Rowher, G.R.** (1984). Scrapie infectious agent is virus like in size and susceptibility to inactivation. *Nature* **308**, 658-662.

**Sambrook, J., Fritsch, E.F., Maniatis, T.** (1989). Molecular cloning: A Laboratory Manual. (Cold Spring Harbour Laboratory., Cold Spring Harbour, NY). 2nd edition.

**Sanger, F., Niklen, S., Coulson, A.R.** (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. (USA).* **74**, 5463-5467.

**Santiago, C., Collins, M., and Johnson, L.F.** (1983). In vitro and in vivo analysis of the control of dihyrofolate reductase gene transcription in serum-stimulated mouse fibroblasts. *J. Cell. Physiol.* **118**, 79-86.

**Scmidt, E.E., Owen, R.A., Merrill, G.** (1990). An intragenic region downstream from the dihydrofolate reductase promoter is required for replication dependent expression. *J. Biol. Chem.* **265** (29), 17397-17400.

**Scott, M., Foster, D., Mirenda, C., Serban, D., Coutal, F., Walchii, M., Torchia, M., Groth, D., Carlson, G., DeArmond, S.J., Westaway, D., and Prusiner, S.B.** (1989). Transgenic mice expressing hamster prion protein produce species specific scrapie infectivity and amyloid plaques. *Cell* **59**, 847-857.

**Seed, B., and Sheen, J.Y.** (1988). A simple phase extraction assay for chloramphenicol acetyltransferase activity. *Gene* **67**, 271-277.

**Smale, S.T., and Baltimore, D.** (1989). The initiator as a transcriptional control element. *Cell* **57**, 103-113.

**Southern, E.M.** (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517.

**Sparkes, R.S., Simon, M., Cohn, V.H., Fournier, R.E.K., Lem, J., Klisak, I., Heinzmann, C., Blatt, C., Lucero, M., Mohandas, T., DeArmond, S.J., Westaway, D., Prusiner, S.B., and Weiner, L.P.** (1986). Assignment of the human and mouse prion protein genes to homologous chromosomes. *Proc. Natl. Acad. Sci. (USA).* **83**, 7358-7362.

**Stahl, N., Borchelt, D.R., Hsiao, K., and Prusiner, S.B.** (1987). Scrapie prion protein contains a phoshatidylinositol glycolipid. *Cell* **51**, 229-240.

**Stahl, N., Borchelt, D.R. and Prusiner, S.B.** (1990). Differential release of cellular and scrapie prion proteins from cellular membranes by phosphatidylinositol-specific phopholipase C. *Biochemistry* **29**, 5405-5412.

**Stahl, N., Baldwin, M.A., Hecker, R., Pan, K.M., Burlingame, A.L., and Prusiner, S.B.** (1992). Glycosylinositol phospholipid anchors of the scrapie and cellular prion proteins contain sialic acid. *Biochemistry* **31**, 5043-5053.

**Stahl, N., Baldwin, M.A., Teplow, D.B., Hood, L., Gibson, B. W., Burlingame, A.L., and Prusiner, S.B.** (1993). Structural analysis of the scrapie prion protein using mass spectrometery and amino acid sequencing. *Biochemistry* **32**, 1991-2002.

**Targa, F.C., De Moura Gallo, C.V., Huesca, M., Scherrer, K., Marcaud, A.W.** (1993). Silencer and enhancer elements located at the 3' side of the chicken and duck α-globin encoding gene domains. *Gene* **129**, 229-237.

**Timmusk, T., Palm, K., Metsis, M., et al.,** (1993). Multiple promoters direct tissue specific expression of rat BDNF gene. *Neuron* **10**, 475-489.

**Turk, E., Teplow, D.B., Hood, L.E., Prusiner, S.B.** (1988). Purification and properties of cellular and scrapie prion protein. *Eur. J. Biochem.* **176**, 21-30.

**Valerio, D., Duyvesteyn, M.C.G., Dekker, B.M.M., Weeda, G., Berkvens, T.M., Van der Voorn, L., Van Ormondt, H., and Van der Eb, A.J.** (1985). Adenosine deamninase: characterisation and expression of a gene with a remarkable promoter. *EMBO. J.* 4 (2), 437-443.

**Wahl, G.M., Lewis, K.A., Ruiz, J.C., Rothenberg, B, Zhao, J., and Evan, G.A.** (1987). Cosmid vectors for rapid genomic walking, restriction mapping and gene transfer. *Proc. Natl. Acad. Sci. (USA).* **84**, 2160-2164.

**Weichenhan, D.** (1991). Fast recovery of DNA from agarose gels by centrifugation through blotting paper. *Trends in Genetics* 7 (4), 109.

**Weissman, C., Bueler, H., Sailer, A., Fischer, M., Auget, M., and Aguzzi, A.** (1993). Role of PrP in prion diseases. *British Medical Bulletin* **49**, (4) 995-1011.

**Wells, G.A.H., Scott, A.C., Johnson, C.T., Gunning, R.F., Hancock, R.D., Jeffrey, M., Dawson, M., Bradley, R.** (1987). A novel progressive spongiform encephalopathy in cattle. *Vet. Rec.* **121**, 419-420.

**Westaway, D., and Prusiner, S.B.** (1986). Conservation of the cellular gene encoding the scrapie prion protein. *Nucleic Acids Research* **14**, 2035-2044.

**Westaway, D., Goodman, P.A., Mirenda, C.A., McKinley, M.P., Carlson, G.A., and Prusiner, S.B.** (1987). Distinct prion proteins in short and long scrapie incubation period mice. *Cell* **51**, 651-662.

**Westaway, D., Mirenda, C.A., Foster, D., Zerbarjadian, Y., Scott, M., Torchia, M., Yang, S.L., Serban, H., DeArmond, S.J., Ebling, C., Prusiner, S.B., and Carlson, G.A.** (1991). Paradoxical shortening of scrapie incubation times by expression of prion protein transgenes derived from long incubation period mice. *Neuron* **7**, 59-68.

**Westaway, D., DeArmond, S.J., Cayetano-Canlas, P., Groth, D., Foster, D., Yang, S.L., Torchia, M., Carlson, G.A., and Prusiner, S.B.** (1994) a. Degeneration of skeletal muscle, peripheral nerves and the central nervous system in transgenic mice overexpressing wild-type prion proteins. *Cell* **76**, 117-129.

**Westaway, D., Zuliani, V., Cooper, C., DaCosta, M., Neuman, S., Jenny A.L., Detwiler, L., Prusiner, S.B.** (1994) b. Homozygosity for prion protein alleles encoding glutamine-171 renders sheep suseceptible to natural scrapie. *Genes and Development* **8**, 959-969.

**Westaway, D., Cooper, C., Turner, S., Da Costa, M., Carlson, G.A., and Prusiner, S.B.** (1994) c. Structure and polymorpism of mouse prion protein gene. *Proc. Natl. Acad. Sci.(USA).* **91**, 6418-6422.

**Wietgrefe, S., Zupancic, M., Haase, A., Chesebro, B., Race, R., Frey, W., Rustan, T., Friedman, R.L.** (1985). Cloning of a gene whose expression is increased in scrapie and senile plaques in human brain. *Science* **230**, 1177-1179.

**Wilesmith J.W, Ryan J.B.M., Atkinson, M.J.** (1991). Bovine spongiform encephalopathy: epidemiological studies on the origin. *Vet. Rec.* **128**, 199-203.

**Wion, D., Le Bert, M., and Brachet, P.** (1988). Messenger RNAs of β-amyloid precursor protein and prion protein are regulated by nerve growth factor in PC12 cells. *Int.J. Dev. Neurosci.* **6**, 387-340.

**Yazdanbakhsh, K., Fraser, P., Kioussis, D., Vidal, M., Grosveld, F., and Lindenbaum, M.** (1993). Functional analysis of the human neurofilament light chain gene promoter. *Nucleic Acids Research* **21**(3), 455-461.

**Young, R.A., Hagenbuchle, O., and Schibler, U.** (1981). A single mouse α-amylase gene specifies two different tissue-specific mRNAs. *Cell* **23**, 451-458.

## APPENDIX 1.

Recipes for chapter 2 solutions.

L-agar: 10 g bactotryptone, 5 g bactoyeast extract, 10 g NaCl, 15 g bactoagar in 1 L $H_2O$ at pH 7.

L-broth: 10 g bactotryptone, 5 g bactoyeast extract, 10 g NaCl in 1 L $H_2O$ at pH 7.

Solution I: 50 mM glucose, 25 mM Tris-HCl pH 8, 10 mM EDTA pH 8.

Solution II: 0.2 M NaOH, 1 % SDS.

Solution III: 3 M Potassium acetate pH 5.

T.E.: 10 mM Tris-HCl pH 8, 1 mM EDTA pH 8.

Cell resuspension solution: 50 mM Tris-HCl pH 7.5, 10 mM EDTA, 100 $\mu g\ ml^{-1}$ RNase.

Cell lysis solution: Same as solution II (above).

Neutralisation solution: 2.55 M Potassium acetate pH 4.8.

Magic mini-prep purification resin: 7 M guanidine HCl.

Wash solution: 200 mM NaCl, 20 mM Tris-HCl pH 7.5, 5 mM EDTA: diluted 1:1 with 95 % ethanol.

1 X TAE: 40 mM Tris-acetate, 1 mM EDTA pH 8.

1 X TBE: 90 mM Tris-borate, 1 mM EDTA pH 8.

Gel loading buffer: 0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF, 15 % (w/v) Ficoll type 400 in $dH_2O$.

10 X CIP dephosphorylation buffer: 500 mM Tris-HCl, 1 mM EDTA pH 8.5.

5 X T4 ligation buffer: 0.25 M Tris-HCl pH 7.6, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25 % (w/v) polyethylene glycol-8000.

Solution A: 470 µl of solution O, 9 µl β mercaptoethanol, 12.5 µl 2 mM dATP, dTTP, dGTP.

Solution O: 1.5 g Tris base, 0.25 g $MgCl_2$ .$6H_2O$ added to 10 ml $H_2O$ at pH 8.

Solution B: 6.5g HEPES (N-2-Hydroxyethyl piperazine-N'-2-ethanesulphonic acid) in 12.5 ml $dH_2O$ pH 6.6.

Solution C: Random hexamer primers $pdN_6$ (Pharmacia) at 50 µg $ml^{-1}$.

10 X kinase buffer: 0.5 M Tris-HCl pH 7.5, 0.1 M $MgCl_2$, 50 mM DTT.

20 X SSC: 175.3 g NaCl, 88.2 g sodium citrate in 1 L $dH_2O$ at pH 7.

Hybond-N transfer buffer: 0.4 M NaOH, 0.6 M NaCl.

Hybond-N (pre) Hybridisation buffer

- random labelled probes: 40 ml 1 M $Na_2PO_4$, 10 ml 1M $NaH_2PO_4$, 50 ml $H_2O$, 7 % SDS, 1 X Denhardts, 5 mg boiled, sonicated and sheared herring sperm DNA.
- oligonucleotide probes: 15 ml 6 X SSC, 0.1 % SDS, 1 X Denhardts, 5 mg boiled, sonicated and sheared herring sperm DNA, 32.5 ml $H_2O$.

Genescreen transfer buffer: 2.3 M NaCl, 2.5 M NaOH.

| | |
|---|---|
| Genescreen (pre) Hybridisation buffer: | 10 ml 50 % dextran sulphate, 15 ml 20 X SSC, 5 ml 10 % SDS, 5 mg boiled, sonicated and sheared herring sperm DNA. |
| Denaturing solution: | 1.5 M NaCl, 0.5 M NaOH. |
| Neutralising solution: | 1.5 M NaCl, 0.5 M Tris-HCl pH 7.2, 1 mM EDTA. |
| Sequenase reaction buffer: | 200 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl. |
| Labelling mix: | 7.5 µM dGTP, 7.5 µM dCTP, 7.5 µM dTTP. |
| Mn buffer: | 0.15 M sodium isocitrate, 0.1 M $MnCl_2$. |
| Stop solution: | 95 % formamide, 20 mM EDTA, 0.05 % (w/v) bromophenol blue, 0.05 % (w/v) xylene cyanol FF. |
| Freezing mix: | 6 ml DMEM, 2 ml DMSO, 2 ml FCS. |

## APPENDIX 2.

Recipes for chapter 3 solutions.

TEN 9: 50 mM Tris-HCl pH 9, 100 mM EDTA, 200 mM NaCl.

SM buffer: 5.8 g NaCl, 2 g $MgSO_4$, 50 ml 1 M Tris-HCl pH 7.5, 5 ml 2 % gelatin.

## APPENDIX 3.

Recipes for chapter 5 solutions.

1 X exonuclease III buffer: 66 mM Tris-HCl pH 8, 0.66 mM $MgCl_2$

7.4 X S1 buffer: 0.3 M potassium acetate pH 4.6, 2.5 M NaCl, 10 mM $ZnSO_4$, 50 % glycerol.

S1 stop buffer: 0.3 M Tris base, 0.05 M EDTA.

HEPES buffer 50 mM HEPES ( [4-(2-hydroxy ethyl)-1-piperazine ethanesulphonic acid), 280 mM $NaCl_2$, 1.5 mM $Na_2HPO_4$, pH 7.1.

Mix A: 46 mM Tris-HCl pH 8, 2.3 µCi D-threo-[dichloracetyl-1-$^{14}$C] chloramphenicol, 231 µM butryl co-A.

100 X Mg buffer: 0.1 M $MgCl_2$, 4.5 M β-mercaptoethanol.

## APPENDIX 4.

Recipes for chapter 6 solutions.

10 X RNA running buffer: 20.92 g MOPS (3-{N-morpholino} propanesulphonic acid), 8.33 M sodium acetate, 10 ml 0.5 M EDTA in 500 ml at pH 7.

Formamide sample buffer: 100 µl 10 X RNA running buffer, 200 µl formamide [99 %], 120 µl formaldehyde [37 40 % w/v].

Ficoll-dye-EDTA: 500 µl 0.2 M EDTA pH 7, 0.3 g Ficoll type 400, trace bromophenol blue.

Loading buffer: 95 % formamide, 20 mM EDTA, 0.05 % (w/v) bromophenol blue, 0.05 % (w/v) xylene cyanol FF.